

STIC-Biotech/ChemLib

92653

From: Mehta, Ashwin
Sent: Monday, April 28, 2003 3:57 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

STIC,

Please search the commercial and interference databases for the nucleotide sequences of SEQ ID NOs: 5 and 6 from 09/810,764.

My mail room is 9E12, office 9E07, art unit 1638.

Thank you,
Ashwin Mehta

Ashwin Mehta
United States Patent and Trademark Office
Biotechnology Patent Examiner
703-306-4540

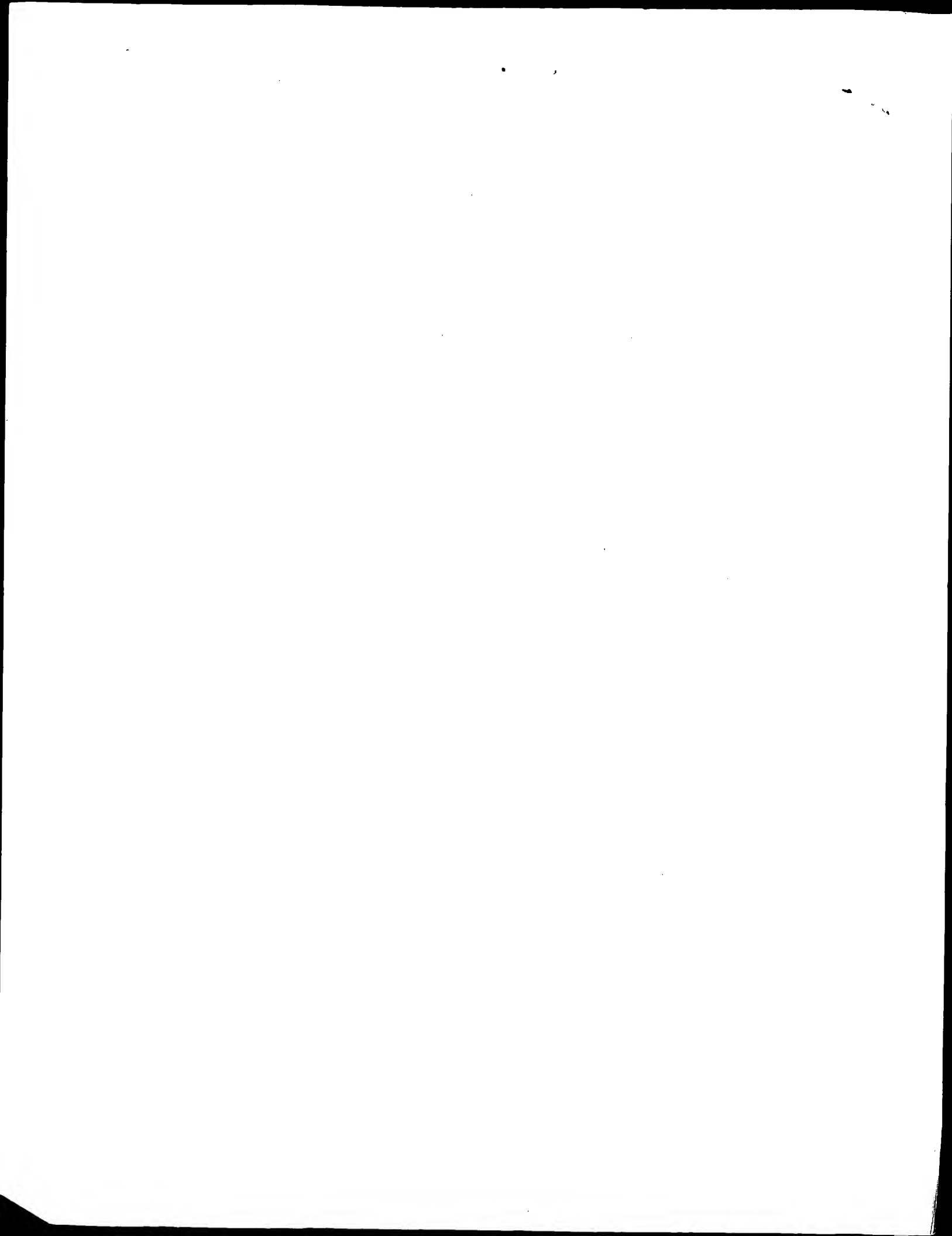
Point of Contact:
Toby Port
Technical Info. Specialist
CM1 6A04
703-308-3534

CR1E

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 4/29
Date Completed: 5/2
Searcher Prep/Review: RV
Clerical: _____
Online time: RV

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:33:59 ; Search time 75 Seconds
(without alignments)
33.757 Million cell updates/sec

Title: US-09-810-764A-5

Perfect score: 90

Sequence: 1 RKGPERKPGVNGTIXXV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	86.7	465	21	AAG40781 Zea mays protein f
2	72	80.0	476	21	AA18102 Arabidopsis thalia
3	72	80.0	476	21	AAG52762 Arabidopsis thalia
4	72	80.0	497	21	AA18101 Arabidopsis thalia
5	72	80.0	497	21	AAG52761 Arabidopsis thalia
6	69	76.7	396	22	AAG93183 C glutamicum prote
7	66	73.3	394	22	AA15890 E. coli proliferat
8	66	73.3	394	22	AAU34769 E. coli cellular p
9	66	73.3	394	22	AAU34838 E. coli cellular p
10	66	73.3	394	22	AAU35466 Haemophilus influe

11	66	73.3	394	22	AAU35476 Haemophilus influe
12	66	73.3	409	22	AAU38371 Salmonella typhi c
13	65	72.2	396	13	AA20242 Sequence of transl
14	65	72.2	396	13	AA20243 Sequence of transl
15	65	72.2	396	13	AA20245 Sequence of elfamy
16	65	72.2	396	13	AA20246 Sequence of elfamy
17	65	72.2	396	13	AA20247 Sequence of elfamy
18	65	72.2	396	13	AA20248 Human structural p
19	63	70.0	397	23	ABP31625 Propionibacterium
20	63	70.0	132	22	AAU66514 Propionibacterium
21	63	70.0	436	22	AAU39992 N-terminal sequenc
22	62	68.9	22	18	AAW33404 Helicobacter pylor
23	62	68.9	22	19	AAW73036 Helicobacter pylor
24	62	68.9	399	19	AAW73035 Helicobacter pylor
25	62	68.9	399	22	AAU35822 Pseudomonas aerugi
26	61	67.8	397	22	AAU36402 Pseudomonas aerugi
27	61	67.8	397	22	AAU36409 Listeria monocytog
28	60	66.7	395	23	AB48242 Elongation factor
29	59	65.6	396	18	AAW30303 Streptococcus pneu
30	57	63.3	398	22	AAU37578 Streptococcus pneu
31	57	63.3	398	23	ABP30281 Streptococcus poly
32	57	63.3	404	23	ABP25922 Streptococcus poly
33	57	63.3	415	23	ABP25923 Drosophila melanog
34	56	62.2	489	22	AB862027 Zea mays protein f
35	55	61.1	77	21	AAG12253 Zea mays protein f
36	55	61.1	86	21	AAG12252 Zea mays protein f
37	55	61.1	111	21	AAG12476 Zea mays protein f
38	55	61.1	129	21	AAG12474 Zea mays protein f
39	55	61.1	156	21	AAG44634 Zea mays protein f
40	55	61.1	175	21	AAG44633 Zea mays protein f
41	55	61.1	426	21	AA03444 Candida albicans e
42	53	58.9	394	20	AAV34672 Chlamydia pneumoni
43	53	58.9	394	22	AAU38889 C. pneumoniae CT32
44	53	58.9	394	22	AAU38909 C. trachomatis CT3
45	53	58.9	394	23	AB94303 Chlamydia pneumoni

ALIGNMENTS

RESULT 1

AAG40781
ID AAG40781 standard; Protein: 465 AA.

XX AAG40781;

AC AAG40781;

XX 18-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 50648.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.

OS Zea mays subsp. mays.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 23-MAR-1999; 99US-0123548.

XX 25-MAR-1999; 99US-0125788.

XX 29-MAR-1999; 99US-0126264.

XX 01-APR-1999; 99US-0127462.

XX 08-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0128714.

XX 19-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.


```

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      86.7%; Score 78; DB 21; Length 465;
Best Local Similarity 84.2%; Pred. No. 5.4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
   ||||| ||||| ||||| |
Db 59 RGFERTKPHVNIQTIGHV 77

RESULT 2
ID AAG18102 standard; Protein; 476 AA.
XX AC AAG18102;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 19377.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

```

PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 80.0%; Score 72; DB 21; Length 476;
 Best Local Similarity 78.9%; Pred. No. 0.00059;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKTKPGVNIIGITXXV 19
 ||||| ||||| |
 Db 70 RGKFKTKPGVNIIGITGHV 88

RESULT 3
 AAG52762
 ID AAG52762 standard; Protein; 476 AA.
 XX
 AC AAG52762;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 67107.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141514.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.


```
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.0%; Score 72; DB 21; Length 497;
Best Local Similarity 78.9%; Pred. No. 0.00062;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKPERTKPGVNIPTXXV 19
   ||||| ||||| |
Db 91 RGKPERKPPHNVNIGTIGHV 109

RESULT 5
AAG52761
ID AAG52761 standard; Protein; 497 AA.
XX
AC AAG52761;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67106.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
```

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.0%; Score 72; DB 21; Length 497;
Best Local Similarity 78.9%; Pred. No. 0.00062;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKTPGVNIGTIXV 19
 Db 91 RGKFKTPGVNIGTIGHV 109

RESULT 6

AAG93183
 ID AAG93183 standard; Protein; 396 AA.

AC AAG93183;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6937.

KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 organic acid synthesis.

OS Corynebacterium glutamicum.

PN EPI108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI: 2001-376931/40.

DR N-PSDB; AAH68402.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT

PS Claim 29; SEQ ID NO: 6937; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium-glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Corynebacterium bacterium, and identifying a homologue of a gene derived
 CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

XX Sequence 396 AA;

Query Match 76.7%; Score 69; DB 22; Length 396;
 Best Local Similarity 73.7%; Pred. No. 0.0016;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKTPGVNIGTIXV 19

Db 3 KAKFKTPGVNIGTIGHV 21

RESULT 7

AAAB15890

ID AAAB15890 standard; Protein; 394 AA.

XX AAAB15890;
 AC 05-OCT-2000 (first entry)
 DT
 DE E. coli proliferation associated protein sequence SEQ ID NO:247.
 XX Escherichia coli; E. coli; proliferation; inhibition; screening;
 KW antimicrobial; bacterial growth; antisense therapy; antibacterial.
 XX Escherichia coli.
 OS
 XX WO2000044906-A2.
 PN 03-AUG-2000.
 PD 27-JAN-2000; 2000WO-US02200.
 PF 27-JAN-1999; 99US-0117405.
 PR (ELIT-) ELITRA PHARM INC.
 PA Zyskind J, Ohlsen KL, Trawick J, Forsyth JM, Froelich JM, Carr GJ;
 PI Yamamoto RT, Xu HH;
 PI WPI: 2000-514822/46.
 DR N-PSDB; AAAB5894.
 DR Novel polynucleotides and polypeptides associated with microorganism
 PT proliferation, used to identify inhibitors of bacterial growth and
 PT proliferation, for use in antisense therapy -
 XX Claim 11; Page 177-178; 316pp; English.
 PS AAAB5809 to AAAB5889 and AAAB66058 to AAAB66138 represent nucleotide
 CC sequences derived from Escherichia coli which inhibit E. coli
 CC proliferation. AAAB5890 to AAAB66055 and AAAB15886 to AAAB16040 represent
 CC nucleotide and protein sequences associated with E. coli proliferation.
 CC AAAB66056 and AAAB66057 represent primers used for sequencing E. coli
 CC proliferation inhibiting nucleotide inserts in an example from the
 CC present invention. Methods from the present invention can be used to
 CC identify a proliferation- required gene in a microorganism, by contacting
 CC a microorganism with a proliferation- required gene activity inhibitory
 CC nucleic acid identified in another organism, and determining if
 CC inhibition occurs in the second microorganism. The nucleic acid sequences
 CC identified as being required for bacterial growth and proliferation, can
 CC be used for antisense therapy for killing bacteria.
 XX Sequence 394 AA;
 SQ
 Query Match 73.3%; Score 66; DB 21; Length 394;
 Best Local Similarity 76.5%; Pred. No. 0.0052;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 KFERTKPGVNIIGTIXV 19
 Db 5 KFERTKPGVNIIGTIGHV 21
 RESULT 8
 AAU34769
 ID AAU34769 standard; Protein; 394 AA.
 XX
 AC AAU34769;
 XX 14-FEB-2002 (first entry)
 DT
 DE E. coli cellular proliferation protein #350.
 XX Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX Escherichia coli.

XX WO200170955-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US09180.
XX 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-206848P.
XX 26-MAY-2000; 2000US-207727P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX N-PSDB; AAS52628.
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 10362; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms. The present sequence represents an
XX essential prokaryotic cellular proliferation protein.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

XX Seq Sequence 394 AA;
XX
XX Query Match 73.3%; Score 66; DB 22; Length 394;
XX Best Local Similarity 76.5%; Pred. No. 0.0052;
XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 3 KFERTKPGVNIQTIXXV 19
XX ||||| ||:| |
XX Db 5 KFERTKPHVNYGTIGHV 21
XX
XX RESULT 9
XX AAU34838
XX ID AAU34838 standard; Protein; 394 AA.
XX
XX AC AAU34838;
XX
XX DT 14-FEB-2002 (first entry)
XX
XX DE E. coli cellular proliferation protein #419.
XX
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX
XX OS Escherichia coli.

XX WO200170955-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US09180.
XX 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-206848P.
XX 26-MAY-2000; 2000US-207727P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX N-PSDB; AAS52697.
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 10431; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms. The present sequence represents an
XX essential prokaryotic cellular proliferation protein.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

XX Seq Sequence 394 AA;
XX
XX Query Match 73.3%; Score 66; DB 22; Length 394;
XX Best Local Similarity 76.5%; Pred. No. 0.0052;
XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 3 KFERTKPGVNIQTIXXV 19
XX ||||| ||:| |
XX Db 5 KFERTKPHVNYGTIGHV 21
XX
XX RESULT 10
XX AAU35466
XX ID AAU35466 standard; Protein; 394 AA.
XX
XX AC AAU35466;
XX
XX DT 14-FEB-2002 (first entry)
XX
XX DE Haemophilus influenzae cellular proliferation protein #107.
XX
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX
XX OS Haemophilus influenzae.

XX WO200170955-A2.
 XX 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US09180.
 XX
 XX 21-MAR-2000; 2000US-191078P.
 XX 23-MAY-2000; 2000US-206848P.
 XX 26-MAY-2000; 2000US-207727P.
 XX 23-OCT-2000; 2000US-242578P.
 XX 27-NOV-2000; 2000US-253625P.
 XX 22-DEC-2000; 2000US-257931P.
 XX 16-FEB-2001; 2001US-269308P.
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX Yamamoto RT, Xu HH;
 XX WPI; 2001-611495/70.
 XX N-PSDB; AAS53325.
 XX
 XX New polynucleotides for the identification and development of
 XX antibiotics, comprise sequences of antisense nucleic acids -
 XX
 XX Example 3; Seq ID No 11059; 511pp; English.
 XX
 XX The invention relates to antisense inhibitors of genes essential to
 XX prokaryotic cellular proliferation, their use in identifying the
 XX genes, themselves and the encoded proteins. The prokaryotes used are
 XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 XX invention is also useful for the identification of potential new targets
 XX for antibiotic development. The antisense nucleic acids can also be used
 XX to identify proteins used in proliferation, to express these proteins,
 XX and to obtain antibodies capable of binding to the expressed proteins.
 XX The proteins can be used to screen compounds in rational drug discovery
 XX programmes. The antisense nucleic acid sequence is also useful to screen
 XX for homologous nucleic acids which are required for cell proliferation in
 XX a wide variety of organisms. The present sequence represents an
 XX essential prokaryotic cellular proliferation protein.
 XX Note: The sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 394 AA;
 XX
 XX Query Match 73.3%; Score 66; DB 22; Length 394;
 XX Best Local Similarity 76.5%; Pred. No. 0.0052;
 XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 XX
 XX QY 3 KFERTKPGVNGTIXXV 19
 XX ||||| ||||| |
 XX Db 5 KFERTKPHVNGTIGHV 21
 XX
 XX RESULT 11
 XX AAU35476
 XX ID AAU35476 standard; Protein; 394 AA.
 XX
 XX AC AAU35476;
 XX
 XX 14-FEB-2002 (first entry)
 XX
 XX Haemophilus influenzae cellular proliferation protein #117.
 XX Antisense; prokaryotic cellular proliferation protein;
 XX antibiotic; antibacterial; drug design.
 XX
 XX Haemophilus influenzae.

XX WO200170955-A2.
 XX 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US09180.
 XX
 XX 21-MAR-2000; 2000US-191078P.
 XX 23-MAY-2000; 2000US-206848P.
 XX 26-MAY-2000; 2000US-207727P.
 XX 23-OCT-2000; 2000US-242578P.
 XX 27-NOV-2000; 2000US-253625P.
 XX 22-DEC-2000; 2000US-257931P.
 XX 16-FEB-2001; 2001US-269308P.
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX Yamamoto RT, Xu HH;
 XX WPI; 2001-611495/70.
 XX N-PSDB; AAS53335.
 XX
 XX New polynucleotides for the identification and development of
 XX antibiotics, comprise sequences of antisense nucleic acids -
 XX
 XX Example 3; Seq ID No 11069; 511pp; English.
 XX
 XX The invention relates to antisense inhibitors of genes essential to
 XX prokaryotic cellular proliferation, their use in identifying the
 XX genes, themselves and the encoded proteins. The prokaryotes used are
 XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 XX invention is also useful for the identification of potential new targets
 XX for antibiotic development. The antisense nucleic acids can also be used
 XX to identify proteins used in proliferation, to express these proteins,
 XX and to obtain antibodies capable of binding to the expressed proteins.
 XX The proteins can be used to screen compounds in rational drug discovery
 XX programmes. The antisense nucleic acid sequence is also useful to screen
 XX for homologous nucleic acids which are required for cell proliferation in
 XX a wide variety of organisms. The present sequence represents an
 XX essential prokaryotic cellular proliferation protein.
 XX Note: The sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 394 AA;
 XX
 XX Query Match 73.3%; Score 66; DB 22; Length 394;
 XX Best Local Similarity 76.5%; Pred. No. 0.0052;
 XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 XX
 XX QY 3 KFERTKPGVNGTIXXV 19
 XX ||||| ||||| |
 XX Db 5 KFERTKPHVNGTIGHV 21
 XX
 XX RESULT 12
 XX AAU38371
 XX ID AAU38371 standard; Protein; 409 AA.
 XX
 XX AC AAU38371;
 XX
 XX 14-FEB-2002 (first entry)
 XX
 XX Salmonella typhi cellular proliferation protein #262.
 XX Antisense; prokaryotic cellular proliferation protein;
 XX antibiotic; antibacterial; drug design.
 XX
 XX Salmonella typhi.

XX WO200170955-A2.
 XX
 XX PD 27-SEP-2001.
 XX
 XX PF 21-MAR-2001; 2001WO-US09180.
 XX
 XX PR 21-MAR-2000; 2000US-191078P.
 XX PR 23-MAY-2000; 2000US-206848P.
 XX PR 26-MAY-2000; 2000US-207727P.
 XX PR 23-OCT-2000; 2000US-242578P.
 XX PR 27-NOV-2000; 2000US-253625P.
 XX PR 22-DEC-2000; 2000US-257931P.
 XX PR 16-FEB-2001; 2001US-269308P.
 XX
 XX PA (ELIT-) ELITRA PHARM INC.
 XX
 XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
 XX PI Yamamoto RT, Xu HH;
 XX
 XX WPI: 2001-611495/70.
 XX DR N-PSDB; AAS56230.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 PT
 XX Example 3; Seq ID No 13964; 511pp; English.
 XX
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 409 AA;
 XX
 XX Query Match 73.3%; Score 66; DB 22; Length 409;
 XX Best Local Similarity 76.5%; Pred. No. 0.0054; 3; Indels 0; Gaps 0;
 XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 KFERTKPGVNGITXXV 19
 Db 20 KFERTKPGVNGITGHV 36
 ||||| |||||
 RESULT 13
 AAR20242
 ID AAR20242 standard; Protein; 396 AA.
 AC AAR20242;
 XX
 XX 15-APR-1992 (first entry)
 DT
 XX Sequence of translation elongation factor Tuf1 encoded by tuf1 gene.
 DE
 XX Elfamycin resistant actinomycetes; antibiotic resistant;
 KW elongation factor.
 XX Streptomyces ramocissimus.
 XX
 XX PI Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 XX PI Woudt LP;
 XX WPI: 1992-017874/03.
 XX DR N-PSDB; AAQ20216.

XX EP466251-A.
 XX
 XX PD 15-JAN-1992.
 XX
 XX PF 02-JUL-1991; 91EP-0201702.
 XX
 XX PR 02-JUL-1991; 91EP-0201702.
 XX PR 10-JUL-1990; 90EP-0201851.
 XX
 XX PA (KONN) GIST-BROCADES NV.
 XX
 XX PI Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 XX PI Woudt LP;
 XX WPI: 1992-017874/03.
 XX DR N-PSDB; AAQ20215.
 XX
 XX New protein conferring resistance to elfamycin - used to
 PT transform streptomycetes to resistant pheno-type
 PT
 XX Example; Fig 1 and Pages 13-15; 35pp; English.
 XX
 XX Substitution of residue 378 of the elongation factor (EF-Tu) with a
 CC valine, threonine, proline or phenylalanine results in an elfamycin
 CC resistant protein (EF-TuR). The advantage of this change is that
 CC the limiting factor for the prodn. of elfamycin by actinomycetes is
 CC removed by mutating the gene tuf into tufR encoding a protein.
 CC resistant to the elfamycin, pref. mocimycin (Kirmomycin). The
 CC inventors claim EF-TuR and the genes (tufR) encoding it.
 XX
 XX Sequence 396 AA;
 XX
 XX Query Match 72.2%; Score 65; DB 13; Length 396;
 XX Best Local Similarity 63.2%; Pred. No. 0.0077; 4; Indels 0; Gaps 0;
 XX Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RGFERTKPGVNGITXXV 19
 Db 2 KAFERTKPHVNMGTIGHI 20
 : ||||| |||||
 RESULT 14
 AAR20243
 ID AAR20243 standard; Protein; 396 AA.
 XX
 XX AAR20243;
 XX
 XX 15-APR-1992 (first entry)
 DT
 XX Sequence of translation elongation factor Tuf2 encoded by tuf2 gene.
 DE
 XX Elfamycin resistant actinomycetes; antibiotic resistant;
 KW elongation factor.
 XX Streptomyces ramocissimus.
 XX
 XX PI Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 XX PI Woudt LP;
 XX WPI: 1992-017874/03.
 XX DR N-PSDB; AAQ20216.

XX New protein conferring resistance to elfamycin - used to
PT transform streptomycetes to resistant pheno-type
XX
XX Example; Pages 16-18; 35pp; English.

CC Substitution of residue 378 of the elongation factor (EF-Tu) with a
CC valine, threonine, proline or phenylalanine results in an elfamycin
CC resistant protein (EF-TuR). The advantage of this change is that
CC the limiting factor for the prodn. of elfamycin by actinomycetes is
CC removed by mutating the gene tuf into tufR encoding a protein,
CC resistant to the elfamycin, pref. mocimycin (Kirromycin). The
CC inventors claim EF-TuR and the genes (tufR) encoding it.

XX Sequence 396 AA;

Query Match 72.2%; Score 65; DB 13; Length 396;
Best Local Similarity 63.2%; Pred. No. 0.0077;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
: ||:|||| ||||| :
Db 2 KAKFQTKPHVNMGTIGHI 20

RESULT 15

AAR20245
ID AAR20245 standard; Protein; 396 AA.

AC AAR20245;

DT 15-APR-1992 (first entry)

XX Sequence of elfamycin-resistant elongation factor EF-TuR Val 378
DE encoded by srtuR1 gene.

XX Elfamycin resistant actinomycetes; antibiotic resistant;
KW elongation factor.

XX Streptomyces ramocissimus.

XX EP466251-A.

XX 15-JAN-1992.

XX 02-JUL-1991; 91EP-0201702.

XX 02-JUL-1991; 91EP-0201702.

XX 10-JUL-1990; 90EP-0201851.

XX (KONN) GIST-BROCADES NV.

XX Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
PI Woudt LP;

XX WPI; 1992-017874/03.

XX N-PSDB; AAQ20218.

XX New protein conferring resistance to elfamycin - used to
PT transform streptomycetes to resistant pheno-type

XX Claim 6; Pages 13-15; 35pp; English.

XX Substitution of residue 378 of the elongation factor (EF-Tu) with a
CC valine, threonine, proline or phenylalanine results in an elfamycin
CC resistant protein (EF-TuR). The advantage of this change is that
CC the limiting factor for the prodn. of elfamycin by actinomycetes is
CC removed by mutating the gene tuf into tufR encoding a protein
CC resistant to elfamycin, pref. mocimycin (Kirromycin). The
CC inventors claim EF-TuR and the genes (tufR) encoding it.

XX Sequence 396 AA;

Query Match 72.2%; Score 65; DB 13; Length 396;
Best Local Similarity 63.2%; Pred. No. 0.0077;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
: ||:|||| ||||| :
Db 2 KAKFQTKPHVNMGTIGHI 20

Search completed: April 29, 2003, 09:40:23
Job time : 77 secs

GenCore version 5.1.4_p5_4578
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:38:19 ; Search time 43 Seconds
(without alignments)
42.478 Million cell updates/sec

Title: US-09-810-764A-5

Perfect score: 90
Sequence: 1 RGKFKTPGVNIGTIXXV 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	82.2	396	D60663	translation elonga
2	72	80.0	409	AB2348	translation elonga
3	72	80.0	476	S09152	translation elonga
4	72	80.0	478	J02240	translation elonga
5	72	80.0	478	S36183	translation elonga
6	72	80.0	479	S21567	translation elonga
7	72	80.0	479	S60659	translation elonga
8	72	80.0	485	S36184	translation elonga
9	70	77.8	18	S39153	translation elonga
10	70	77.8	46	S23851	translation elonga
11	70	77.8	409	S78248	translation elonga
12	69	76.7	391	AE3346	protein translatio
13	69	76.7	394	T71672	translation elonga
14	69	76.7	394	S62726	translation elonga
15	69	76.7	394	H97825	elongation factor
16	69	76.7	396	S31151	translation elonga
17	69	76.7	396	G87143	elongation factor
18	69	76.7	405	E75533	translation elonga
19	69	76.7	406	AH3344	translation elonga
20	69	76.7	409	EFEGT	protein translatio
21	69	76.7	488	T06821	translation elonga
22	68	75.6	397	S23908	translation elonga
23	68	75.6	397	S50138	translation elonga
24	68	75.6	397	PC4060	translation elonga
25	68	75.6	405	CG0663	translation elonga
26	68	75.6	409	S04430	translation elonga
27	67	74.4	394	B60663	translation elonga
28	67	74.4	397	H97285	elongation factor
29	67	74.4	410	S04391	translation elonga

30 66 73.3 394 1 ERECT
31 66 73.3 394 1 ERECTA
32 66 73.3 394 2 S13561
33 66 73.3 394 2 S13560
34 66 73.3 394 2 E64078
35 66 73.3 394 2 F91152
36 66 73.3 394 2 G91241
37 66 73.3 394 2 D82332
38 66 73.3 394 2 G82337
39 66 73.3 394 2 D86089
40 66 73.3 394 2 AB0457
41 66 73.3 394 2 AE0025
42 66 73.3 394 2 B85998
43 66 73.3 394 2 AD0934
44 66 73.3 394 2 AD1005
45 66 73.3 396 2 A44795

ALIGNMENTS

RESULT 1

D60663

translation elongation factor EF-Tu - Pseudomonas cepacia

C:Species: Pseudomonas cepacia

C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 02-Feb-2001

C:Accession: D60663

R:Ludwig, W.; Weizenegger, M.; Betzl, D.; Leidel, E.; Lenz, T.; Ludvigsen, A.; Moell

Arch. Microbiol. 153, 241-247, 1990

A:Title: Complete nucleotide sequences of seven eubacterial genes coding for the elongation factor Tu

A:Reference number: A60663; MUID:90240875; PMID:2110445

A:Accession: D60663

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-396 <IUD>

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu

C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis

F:13-139/Domain: translation elongation factor Tu homology <ETU>

F:19-26/Region: nucleotide-binding motif A (P-loop)

F:136-139/Region: GTP-binding NKXD motif

F:174-176/Region: GTP-binding SAK/L motif

F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asp, Ser)

Query Match

Best Local Similarity 82.2%; Score 74; DB 2; Length 396;

Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFKTPGVNIGTIXXV 19

DB 3 GKGFERTKPHVNVGIGHV 21

RESULT 2

AB2348

translation elongation factor EF-Tu [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AB2348

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigun

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB2348

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-409 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAH76036.1; PID:g17133473; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: tufa

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu

us-09-810-764a-5.rpr

Fri May 2 08:51:22 2003

C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis
 F:1-69/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:70-478/Product: translation elongation factor Tu #status predicted <MAT>
 F:82-208/Domain: translation elongation factor Tu homology <ETU>
 F:88-95/Region: nucleotide-binding motif A (P-loop)
 F:205-208/Region: GTP-binding NKXD motif
 F:243-245/Region: GTP-binding SAK/L motif
 F:94,95,131,205,206,208,243/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser)

Query Match 80.0%; Score 72; DB 2; Length 478;
 Best Local Similarity 78.9%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 | ||||| || ||||| |
 Db 72 RGFERTKPGVNIQTIXV 90

RESULT 3
 S36183
 translation elongation factor EF-Tu.A precursor, chloroplast - wood tobacco
 C:Species: Nicotiana sylvestris (wood tobacco)
 C:Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 02-Feb-2001
 C:Accession: S36183; S37674; S46548; S40614
 R:Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiyama, M.
 Plant Mol. Biol. 22, 767-774, 1993
 A:Title: Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco
 A:Reference number: S36183; MUID:93363910; PMID:8958028
 A:Accession: S36183
 A:Molecule type: mRNA
 A:Residues: 1-478 <MUR1>
 A:Cross-references: GB:D11375
 R:Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiyama, M.
 submitted to the EMBL Data Library, June 1992
 A:Reference number: S37674
 A:Accession: S37674
 A:Molecule type: mRNA
 A:Residues: 1-457 <MUR2>
 A:Cross-references: NID:g218309; PIDN:BAA01974.1; PID:g218310
 R:Sugita, M.; Murayama, Y.; Sugiyama, M.
 submitted to the EMBL Data Library, June 1992
 A:Description: Structure and differential expression of two distinct genes encoding
 A:Reference number: S46548
 A:Accession: S46548
 A:Molecule type: DNA
 A:Residues: 1-478 <SUG1>
 A:Cross-references: EMBL:D11469; NID:g459238; PIDN:BAA02027.1; PID:g459239
 R:Sugita, M.; Murayama, Y.; Sugiyama, M.
 Curr. Genet. 25, 164-168, 1994
 A:Title: Structure and differential expression of two distinct genes encoding the ch
 A:Reference number: S40614; MUID:94373864; PMID:8087886
 A:Accession: S40614
 A:Molecule type: DNA
 A:Residues: 1-111 <SUG2>
 A:Cross-references: EMBL:D11469
 C:Genetics:
 A:Gene: tufa
 C:Superfamily: translation elongation factor Tu; translation elongation factor Tu ho
 C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis
 F:1-70/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:71-478/Product: translation elongation factor Tu.A #status predicted <MAT>
 F:82-208/Domain: translation elongation factor Tu homology <ETU>
 F:88-95/Region: nucleotide-binding motif A (P-loop)
 F:205-208/Region: GTP-binding NKXD motif
 F:243-245/Region: GTP-binding SAK/L motif
 F:94,95,131,205,206,208,243/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser)

Query Match 80.0%; Score 72; DB 2; Length 478;
 Best Local Similarity 78.9%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 | ||||| || ||||| |

us-09-810-764a-5.rpr

C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis
 F:1-69/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:70-478/Product: translation elongation factor Tu #status predicted <MAT>
 F:82-208/Domain: translation elongation factor Tu homology <ETU>
 F:88-95/Region: nucleotide-binding motif A (P-loop)
 F:205-208/Region: GTP-binding NKXD motif
 F:243-245/Region: GTP-binding SAK/L motif
 F:94,95,131,205,206,208,243/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser)

Query Match 80.0%; Score 72; DB 2; Length 409;
 Best Local Similarity 78.9%; Pred. No. 0.00021;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 | ||||| || ||||| |
 Db 3 RGFERTKPGVNIQTIXV 21

RESULT 3
 S09152
 translation elongation factor EF-Tu precursor, chloroplast - Arabidopsis thaliana
 N:Alternate names: protein F9f13.10
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 02-Feb-2001
 C:Accession: S09152; T10578
 R:Baldauf, S.L.; Palmer, J.D.
 Nature 344, 262-265, 1990
 A:Title: Evolutionary transfer of the chloroplast tufa gene to the nucleus.
 A:Reference number: S09152; MUID:90190846; PMID:2314461
 A:Accession: S09152
 A:Molecule type: DNA
 A:Residues: 1-476 <BAL>
 A:Cross-references: EMBL:X52256; NID:g22564; PIDN:CA36498.1; PID:g22565
 R:Bevan, M.; Pohl, T.; Weissenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16991
 A:Accession: T10578
 A:Molecule type: DNA
 A:Residues: 15-476 <BEV>
 A:Cross-references: EMBL:AL080253; GSPDB:GN00062; ATSP:F9f13.10
 A:Experimental source: Cultivar Columbia; BAC clone F9f13
 C:Genetics:
 A:Gene: tufa; ATSP:F9f13.10
 A:Map position: 4
 A:Genome: nuclear
 C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homol
 C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis
 F:1-67/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:68-476/Product: translation elongation factor Tu #status predicted <MAT>
 F:80-206/Domain: translation elongation factor Tu homology <ETU>
 F:86-93/Region: nucleotide-binding motif A (P-loop)
 F:203-206/Region: GTP-binding NKXD motif
 F:241-243/Region: GTP-binding SAK/L motif
 F:92,93,129,203,204,206,241/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #st

Query Match 80.0%; Score 72; DB 2; Length 476;
 Best Local Similarity 78.9%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
 | ||||| || ||||| |
 Db 70 RGFERTKPGVNIQTIXV 88

RESULT 4
 JQ2240
 translation elongation factor EF-Tu precursor - common tobacco chloroplast
 C:Species: chloroplast Nicotiana tabacum (common tobacco)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 02-Feb-2001
 C:Accession: JQ2240
 R:Ursin, V.M.; Becker, C.K.; Shewmaker, C.K.
 Plant Physiol. 101, 333-334, 1993
 A:Title: Cloning and nucleotide sequence of a tobacco chloroplast translational elongat
 A:Reference number: JQ2240; MUID:94105295; PMID:8278900
 A:Accession: JQ2240
 A:Molecule type: DNA
 A:Residues: 1-478 <URS>
 C:Comment: This protein is an essential component for protein synthesis that functions b
 C:Genetics:
 A:Genome: chloroplast
 C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homol

Db 72 RGKFERKPPHVNIGTIGHV 90

RESULT 6

S21567

translation elongation factor EF-Tu precursor - soybean chloroplast

C:Species: Chloroplast Glycine max (soybean)

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001

C:Accession: S21567

R:Bonny C.; Stutz, E.

submitted to the EMBL Data Library, May 1992

A:Description: Soybean (glycine max l.) nuclear DNA: structure and expression of chlorop

A:Reference number: S21567

A:Accession: S21567

A:Molecule type: DNA

A:Residues: 1-479 <HON>

A:Cross-references: EMBL:X66062; NID:gl8775; PIDN:CAA46864.1; PID:gl8776

C:Genetics:

A:Gene: tuftA

A:Genome: chloroplast

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo

C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis

F:83-209/Domain: translation elongation factor Tu homology <ETU>

F:89-96/Region: nucleotide-binding motif A (P-loop)

F:206-209/Region: GTP-binding NKXD motif

F:244-246/Region: GTP-binding SAK/L motif

F:95,96,132,206,207,209,244/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #st

Query Match

Best Local Similarity 80.0%; Score 72; DB 2; Length 479;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFERKPPHVNIGTIGHV 91

Db 73 RGKFERKPPHVNIGTIGHV 91

RESULT 7

S60659

translation elongation factor EF-Tu precursor, chloroplast - soybean

C:Species: Glycine max (soybean)

C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 02-Feb-2001

C:Accession: S60659

R:Haurer, F.; Murne, M.; Stutz, E.

Submitted to the EMBL Data Library, July 1995

A:Description: The nuclear genome of soybean contains four tuf genes belonging to two su

A:Reference number: S60659

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-479 <MAU>

A:Cross-references: EMBL:X89058; NID:g949872; PIDN:CAA61444.1; PID:g949873

C:Genetics:

A:Genome: nuclear

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo

C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis

F:83-209/Domain: translation elongation factor Tu homology <ETU>

F:89-96/Region: nucleotide-binding motif A (P-loop)

F:206-209/Region: GTP-binding NKXD motif

F:244-246/Region: GTP-binding SAK/L motif

F:95,96,132,206,207,209,244/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #st

Query Match

Best Local Similarity 80.0%; Score 72; DB 2; Length 479;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFERKPPHVNIGTIGHV 91

Db 73 RGKFERKPPHVNIGTIGHV 91

RESULT 8

S36184

translation elongation factor EF-Tu.B precursor, chloroplast - wood tobacco

C:Species: Nicotiana sylvestris (wood tobacco)

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 02-Feb-2001

C:Accession: S36184; S37675; S46549; S40615

R:Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiura, M.

Plant Mol. Biol. 22, 767-774, 1993

A:Title: Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco

A:Reference number: S36183; MUID:93363910; PMID:8358028

A:Accession: S36184

A:Molecule type: mRNA

A:Residues: 1-485 <MUR>

A:Cross-references: GB:D11376

R:Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiura, M.

submitted to the EMBL Data Library, June 1992

A:Reference number: S37674

A:Accession: S37675

A:Molecule type: mRNA

A:Residues: 61-430, 'IMS', 434-448, 'NM', 451-482, 'ILE', 485 <MU2>

A:Cross-references: GB:D11376; NID:g218311; PIDN:BAA01975.1; PID:g218312

R:Sugita, M.; Murayama, Y.; Sugiura, M.

submitted to the EMBL Data Library, June 1992

A:Description: Structure and differential expression of two distinct genes encoding

A:Reference number: S46548

A:Accession: S46549

A:Molecule type: mRNA

A:Residues: 1-430, 'IMS', 434-448, 'NM', 451-482, 'IL', 485 <SUG1>

A:Cross-references: EMBL:D11470; NID:g459240; PIDN:BAA02028.1; PID:g459241

R:Sugita, M.; Murayama, Y.; Sugiura, M.

Curr. Genet. 25, 164-168, 1994

A:Title: Structure and differential expression of two distinct genes encoding the ch

A:Reference number: S40614; MUID:94373864; PMID:8087886

A:Accession: S40615

A:Molecule type: DNA

A:Residues: 1-103 <SUG2>

C:Genetics:

A:Gene: tufB

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthes

F:1-77/Domain: transit peptide (chloroplast) #status predicted <NP>

F:78-485/Product: translation elongation factor Tu.B #status predicted <MAT>

F:89-215/Domain: translation elongation factor Tu.B #status predicted <MAT>

F:95-102/Region: nucleotide-binding motif A (P-loop)

F:212-215/Region: GTP-binding NKXD motif

F:250-252/Region: GTP-binding SAK/L motif

F:101,102,138,212,213,215,250/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser)

Query Match

Best Local Similarity 80.0%; Score 72; DB 2; Length 485;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFERKPPHVNIGTIGHV 19

Db 79 RGKFERKPPHVNIGTIGHV 97

RESULT 9

S39153

translation elongation factor EF-Tu, chloroplast - common tobacco (fragment)

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 05-Dec-1997

C:Accession: S39153

R:Murayama, Y.; Matsubayashi, T.; Sugita, M.; Sugiura, M.

Plant Mol. Biol. 22, 767-774, 1993

A:Title: Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco

A:Reference number: S36183; MUID:93363910; PMID:8358028

A:Accession: S39153

A:Molecule type: protein

A:Residues: 1-18 <MUR>

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

C:Keywords: chloroplast; GTP binding; protein biosynthesis

Query Match

Best Local Similarity 77.8%; Score 70; DB 2; Length 18;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 16
| | | | | | | | | |
Db 2 RGFERTKPGVNIQT 17

RESULT 10

S23851
translation elongation factor EF-Tu - Eikenella corrodens (fragment)
C:Species: Eikenella corrodens
C:Date: 12-Feb-1993 #sequence_revision 08-Nov-1996 #text_change 02-Feb-2001
C:Accession: S23851
R: Rao, V.K.; Whitlock, J.A.; Progulskie-Fox, A.
submitted to the EMBL Data Library, June 1992
A:Description: Cloning, characterization, and sequencing of two hemagglutinin genes from
A:Reference number: S23847
A:Accession: S23851
A:Molecule type: DNA
A:Residues: 1-46 <RAO>
A:Cross-references: EMBL:Z12610; NID:g41653; PIDN:CAA78258.1; PID:g41658
A:Experimental source: Eikenella corrodens ATCC 23834
C:Genetics:
A:Gene: tuftA
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog
C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis
F:19-26/Region: nucleotide-binding motif A (P-loop)

Query Match 77.8%; Score 70; DB 2; Length 46;

Best Local Similarity 68.4%; Pred. No. 5e-05;

Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 19
: | | | | | | | | | |
Db 3 RGFERTKPGVNIQT 21

RESULT 11

S78248
translation elongation factor EF-Tu - Odontella sinensis chloroplast
C:Species: chloroplast Odontella sinensis
C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 02-Feb-2001
C:Accession: S78248
R: Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
Plant Mol. Biol. Rep. 13, 336-342, 1995
A:Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sinensis
A:Reference number: S78238
A:Accession: S78248
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-409 <KOW>
A:Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91621.1; PID:g1185138
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
C:Genetics:
A:Gene: tuftA
A:Genome: Chloroplast
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog
C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis
F:13-139/Domain: translation elongation factor Tu homology <ETU>
F:19-26/Region: nucleotide-binding motif A (P-loop)
F:136-139/Region: GTP-binding NKXD motif
F:174-176/Region: GTP-binding SAK/L motif
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 77.8%; Score 70; DB 2; Length 409;

Best Local Similarity 78.9%; Pred. No. 0.00046;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 19
| | | | | | | | | |
Db 3 RGFERTKPGVNIQT 21

translation elongation factor Tu (EF-tu) [imported] - Brucella melitensis (st
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AE3346
R: DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3346
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL51936.1; PID:g17982694; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0755
A:Map position: 1
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match 76.7%; Score 69; DB 2; Length 391;

Best Local Similarity 73.7%; Pred. No. 0.00065;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 19
: | | | | | | | | | |
Db 3 RGFERTKPGVNIQT 21

translation elongation factor tu (EF-tu) UU522 [similarity] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 02-Feb-2001
C:Accession: S62726; F62879
R: Kamla, V.

Query Match 76.7%; Score 69; DB 2; Length 394;

Best Local Similarity 73.7%; Pred. No. 0.00066;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQT 19
| | | | | | | | | |
Db 3 RGFERTKPGVNIQT 21

submitted to the EMBL Data Library, June 1994

A:Reference number: S62726
 A:Accession: S62726
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-394 <RAM>
 A:Cross-references: EMBL:Z34275; NID:g498790; PIDN:CAA84029.1; PID:g498791
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to GenBank, February 2000
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mir
 A:Reference number: A82870
 A:Accession: F82879
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-394 <GLA>
 A:Cross-references: GB:AE002151; GB:AF222894; NID:g6899524; PIDN:AAF30935.1; GSPDB:GN001
 A:Experimental source: serovar 3; biovar 1
 C:Genetics:
 A:Gene: tuf; UU522
 A:Genetic code: SGC3
 C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:13-139/Domain: translation elongation factor Tu homology <ETU>
 F:19-26/Region: nucleotide-binding motif A (P-loop)
 F:136-139/Region: GTP-binding NKXD motif

Query Match 76.7%; Score 69; DB 2; Length 394;

Best Local Similarity 73.7%; Pred. No. 0.00066;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGRFERTKPGVNIQTIXV 19

: ||||| |||||

Db 3 KAKFERTKPHVNIQTICHV 21

RESULT 15

H97825

elongation factor EF-Tu [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001

C:Accession: H97825

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; RC
 Science 293; 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: H97825

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-394 <KUR>

A:Cross-references: GB:AE006914; PIDN:AAL03546.1; PID:g15620123; GSPDB:GN00173

C:Genetics:

A:Gene: tuf

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo

Query Match

Best Local Similarity 76.7%; Score 69; DB 2; Length 394;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGRFERTKPGVNIQTIXV 19

: ||||| |||||

Db 3 KAKFERTKPHVNIQTICHV 21

Search completed: April 29, 2003, 09:42:20

Job time : 44 secs

OM protein - protein search, using sw model
Run on: April 29, 2003, 09:36:24 ; Search time 25 Seconds
(without alignments)
31.522 Million cell updates/sec

Title: US-09-810-764A-5
Perfect score: 90
Sequence: 1 RGPRTKPGVNIQTIXV 19
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	83.3	391	1 EFTU_RHILO	098167 rhizobium l
2	74	82.2	396	1 EFTU_BURCE	P31677 burkholderi
3	72	80.0	476	1 EFTU_ARATH	P17745 arabidopsis
4	72	80.0	478	1 EFTU_TORAC	P41342 nicotiana t
5	72	80.0	479	1 EFTU_TOYBN	Q43467 glycine max
6	72	80.0	479	1 EFTU_TOYBN	P35644 glycine max
7	70	77.8	46	1 EFTU_EIKCO	Q94622 odontella s
8	70	77.8	409	1 EFTU_ODOSI	P48865 rickettsia
9	70	77.8	410	1 EFTU_NEPOL	P50068 ureaplasma
10	69	76.7	394	1 EFTU_RICPR	O50306 bacillus st
11	69	76.7	395	1 EFTU_UREPA	P42439 corynebacte
12	69	76.7	395	1 EFTU_RACST	P30768 mycobacteri
13	69	76.7	396	1 EFTU_MYCLE	P95724 streptomyce
14	69	76.7	396	1 EFTU_CORGL	O93342 deinoxococc
15	69	76.7	397	1 EFTU_STRCJ	P02991 euglena gra
16	69	76.7	405	1 EFTU_DEIRA	O24310 plasmu sativ
17	69	76.7	405	1 EFTU_STRCJ	P40174 thioabacillu
18	69	76.7	488	1 EFTU_PEA	Q43110 streptomyce
19	68	75.6	396	1 EFTU_THICO	Q38171 streptomyce
20	68	75.6	397	1 EFTU_STRCO	P29542 streptomyce
21	68	75.6	397	1 EFTI_STRCU	O33594 streptomyce
22	68	75.6	397	1 EFTI_STRCU	P33168 deinonema s
23	68	75.6	397	1 EFTI_STRCU	P18668 synchococc
24	68	75.6	405	1 EFTU_STRAU	P33171 synchococc
25	68	75.6	409	1 EFTU_DEISP	P19457 guillardia
26	68	75.6	409	1 EFTU_SINP6	P13552 spirulina p
27	67	74.4	409	1 EFTU_SINP7	P02990 escherichia
28	67	74.4	408	1 EFTU_BACFR	P43926 haemophilus
29	67	74.4	408	1 EFTU_GUITH	P21694 salmonella
30	66	73.3	393	1 EFTU_SPLI	P57939 pasteurella
31	66	73.3	393	1 EFTU_SPLI	
32	66	73.3	393	1 EFTU_ECOLI	
33	66	73.3	393	1 EFTU_HAEIN	
34	66	73.3	394	1 EFTU_SALTY	
35	66	73.3	394	1 EFTU_PASMU	

ALIGNMENTS

RESULT 1
EFTU_RHILO
ID EFTU_RHILO STANDARD; PRT; 391 AA.
AC Q981F7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Elongation factor Tu (EF-Tu)
GN (TUPA OR MLR0263) AND (TUFB OR MLR0288).
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Watanabe A., Igesawa K., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Kishida Y., Kiyokawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Mochizuki Y., Nakayama S., Kobara M., Matsumoto M., Matsuno A.,
RA Takeuchi C., Yamada M., Nakazaki N., Shimpo S., Sugimoto M.,
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti";
RL DNA Res 7:331-338(2004).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP002994; BAB47886.1;
CC EMBL; AP002994; BAB47904.1;
CC InterPro; IPR004541; EF-Tu.
CC InterPro; IPR004160; EFTU_Cterm.
CC InterPro; IPR004161; EFTU_D2.
CC InterPro; IPR000795; EF-GTPbind.
CC Pfam; PF00009; Small GTP.
CC Pfam; PF03143; GTP_EFTU; 1.
CC PRINTS; PR00315; ELONGATNEC.
CC TIGRfam; TIGR00231; small_GTP; 1.
CC TIGRfam; TIGR00485; EF-Tu; 1.
CC PROSITE; PS00301; EFACOR_GTP; 1.
CC Elongation factor; Protein biosynthesis; GTP-binding;
KW

P57966 pasteurella
P31501 mycobacteri
P51287 porphyra pu
P17746 chlamydomon
P29543 streptomyce
P75022 agrobacteri
P42471 brevibacter
P74227 synchocyst
P07157 thermus the
P14634 astasia lon
Q9mup0 mesostigma
P48864 neisseria g

Fri May 2 08:51:22 2003

RESULT 3

ID	EFTU_ARATH	STANDARD:	PRT:	476 AA.
AC	01-AUG-1990 (Rel. 15, Last sequence update)			
AT	01-AUG-1990 (Rel. 15, Last annotation update)			
CT	01-AUG-1990 (Rel. 40, Last annotation precursor (EF-Tu))			
DT	16-OCT-2001 (Rel. 40, Last annotation precursor (EF-Tu))			
DE	TUFA OR AT4G20360 OR P9f13.10			
DN	Arabidopsis thaliana (Mouse ear cross)			
OS	Arabiopsida: Viridiplantae; Streptophyta; Embryophyta: Tracheophyta;			
OC	Spermatophyta: Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids			
OX	NCBI_TaxID=3702;			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=cv. Landsberg erecta;			
RC	MEDLINE=90190846; PubMed=2314461;			
RX	Balduf S.L., Palmer J.D.;			
RA	"Evolutionary transfer of the chloroplast tuft gene to the nucleus."			
RT	Nature 344:262-265(1990).			
RL	(1)			
CC	SEQUENCE FROM N.A.			
CC	STRAIN=cv. Columbia;			
CC	MEDLINE=20083488; PubMed=10617198;			
CC	Mayer K.F.X., Schuelter A., Striekema W., Entian K.-D., Terryn N.,			
CC	Pohl T., Duesterhoeft A., Brandt P., Grivell L.A., Mache R., Mueller M.,			
CC	Harris B., Ansorge W., Simone V., Obermaier B.M., Schmidheini F.I.,			
CC	Weichselgartner M., de Simone V., Puigdomenech P., Watson M., Bancroft I.,			
CC	Kreis M., Delsen M., Portetelie D., Perez-Alonso M., Bouty M., Ridley P.,			
CC	Reichert B., portetelie D., Zimmermann W., Wedler H.,			
CC	Vos P., Hoheisel J., Zimmermann W., Bilham L., Robben J., Vandebussche F.,			
CC	Langham S.-A., McCullagh B., Voeltz G., Bastiaens I., Aert R., Braun M.,			
CC	Van der Schueren J., Grymonprez B., Voeltz G., Bastiaens I., Aert R., Braun M.,			
CC	Braeken T., Bothe G., Kampinger U., Hilbert H., Dirks W.,			
CC	Weitenecker T., Peters R., Rose M., Hauf J., Koetter P.,			
CC	Holzner E., Brandt A., Klein Lankhorst Rausch M., Lanberth S., van den Daele H.,			
CC	Moolijman P., Hempel S., Feldman R., Villarroel R., de Clercq R.,			
CC	Berneiser A., Buyschaert C., Gielen J., Quail M., Bray-Allen S., Mayes R.,			
CC	De Keyser A., Rogers J., Cronin A., Quail M., Lennard N., McClay K.,			
CC	van Montagu M., Rogers J., Hall S., Kay M., Benes V., Rehnert T.H.,			
CC	Clark L., Doggett J., Hall S., Kay M., Benes V., Rehnert T.H.,			
CC	Pettett A., Rajandream M.A., Lyne M., Grimm M., Loewner Auer S.,			
CC	Borkova D., Bloeker H., Scharfe A.C., Schaefer K., Dauner D., Herzl A.,			
CC	Dose S., de Haan M., Maare A.C., Schaefer K., Dauner D., Herzl A.,			
CC	Gabel C., Fuchs M., Partmann B., Grandrath K., Dauner D., Herzl A.,			
CC	Neumann S., Argüelles A., Vitale D., Ligouri R., Piravandi E.,			
CC	Massenet O., Quigley F., Clabaud G., Muendlein A., Feiler R.,			
CC	Schnabl F., Cooke R., Berger C., Monfort A., Casacuberta E.,			
CC	Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Jesse T.,			
CC	Perez-Perez A., Furnelle B., Bent E., Johnson S., Tacon D., Bielke C.,			
CC	Friehnen L., Haase D., Lemcke K., Mewes H.-W., Stocker S., Habermann K.,			
CC	Zaccaria P., Hevan M., Wilson R.K., de la Bastide M., Habermann K.,			
CC	Parnell L., Dedhia N., Gnoj L., Schulz K., Huang E., Spiegel L.,			
CC	Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,			
CC	Stonking T., Kalicki J., Graves T., Harmon G., Edwards J.,			
CC	Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,			
CC	Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,			
CC	Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier D.,			
CC	Nelson J., Spieth J., Ryan A., Andrews S., K. Cotton M., Joshua C.,			
CC	Du H., Ali J., Bernhoff A., Jones K., Drone K., Lanar B., Jordan C.,			
CC	Anatoliu B., Zidanik M., Strong C., Sun H., Matero A., Shah R.,			
CC	Ma P., Zhong J., Preston A., Rodriguez M., Hoffman J., Till S., A.,			
CC	Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., A.,			
CC	Granat S., Shohdy N., Hasagawa A., Hameed A.,			
CC	Chen E., Marra M., Martienssen R., McCombie W.R.;			
CC	"Sequence and analysis of chromosome 4 of the plant Arabidopsis			
CC	thaliana."			
CC	Nature 402:769-777(1999).			
CC	-1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF			
CC	AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN			

Query Match 82.2% Score 74; DB 1; Length 396;
Best Local Similarity 73.7% Pred. No. 1;e-05; Indels
Matches 14; Conservative 2; Mismatches 3; Gaps 0;

RESULT 2

ID	EFTU_BURCE	STANDARD:	PRT:	396 AA.
AC	P33167; 1993 (Rel. 27, Created)			
AT	01-OCT-1993 (Rel. 27, Last sequence update)			
CT	01-OCT-1993 (Rel. 40, Last annotation update)			</

CC BIOSYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-TU/EF-1A SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X52256; CA336498.1;
 CC EMBL; AL080253; CAB45802.2;
 CC EMBL; AL161552; CAB79036.1;
 CC PIR; S09152; S09152.
 CC HSP; P02990; LEFU.
 CC SWISS-2DPAGE; P17745; ARATH.
 CC InterPro; IPR004541; EF-Tu.
 CC InterPro; IPR004160; EFTU_Cterm.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR000795; EF-GTPbind.
 CC Pfam; PF00009; Small_GTP.
 CC Pfam; PF03143; GTP_EFTU_D3; 1.
 CC Pfam; PF03144; GTP_EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFC.
 CC TIGRFAMS; TIGR00231; small_GTP; 1.
 CC TIGRFAMS; TIGR00485; EF-Tu; 1.
 CC PROSITE; PS00301; EFATOR_GTP; 1.
 CC Elongation factor; Protein biosynthesis; GTP-binding;
 CC Transit peptide; Chloroplast.
 CC TRANSIT
 CC CHAIN 1 67 CHLOROPLAST (POTENTIAL).
 CC FT NP_BIND 68 476 ELONGATION FACTOR TU.
 CC FT NP_BIND 86 93 GTP (BY SIMILARITY).
 CC FT NP_BIND 148 152 GTP (BY SIMILARITY).
 CC FT NP_BIND 203 206 GTP (BY SIMILARITY).
 CC SQ SEQUENCE 476 AA; 51630 MW; 6041F540224A1738 CRC64;
 CC
 CC Query Match 80.0%; Score 72; DB 1; Length 476;
 CC Best Local Similarity 78.9%; Pred. No. 4.3e-05;
 CC Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 1 RKGFEKTKPGVNTGTXV 19
 CC ||||| || ||||| |
 CC Db 70 RKGFEKTKPGVNTGTXV 88
 CC
 CC RESULT 4
 CC EFTU_TOBAC STANDARD; PRT; 478 AA.
 CC ID EFTU_TOBAC
 CC AC P41342;
 CC DT 01-FEB-1995 (Rel. 31, Created)
 CC DT 01-FEB-1995 (Rel. 31, Last sequence update)
 CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
 CC DE Elongation factor Tu, chloroplast precursor (EF-Tu).
 CC GN TUFA.
 CC OS Nicotiana tabacum (Common tobacco), and
 CC OS Nicotiana sylvestris (Wood tobacco).
 CC OC Eukaryota; Viridiplantae; Streptophyta;
 CC OC Spermatophyta; Magnoliophyta; Embryophyta; Tracheophyta;
 CC OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 CC OX NCBI_TaxID=4097, 4096;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC SPECIES-N.Tabacum; STRAIN=cv. NK 236;
 CC RA MEDLINE=94105295; PubMed=8278500;
 CC RX Ursin V.M., Becker C.K., Shewmaker C.K.;
 CC RT "Cloning and nucleotide sequence of a tobacco chloroplast
 CC translational elongation factor, EF-Tu*;
 CC RL plant Physiol. 101:333-334(1993).
 CC [2]

RP SEQUENCE FROM N.A.
 RC SPECIES-N.Sylvestris;
 RX MEDLINE=94373864; PubMed=8087886;
 RA Sugita M., Murayama Y., Sugiura M.;
 RT "Structure and differential expression of two distinct genes encoding
 RT the chloroplast elongation factor Tu in tobacco*;
 RL Curr. Genet. 25:164-168(1994).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
 CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
 CC BIOSYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-TU/EF-1A SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M94204; AAA18546.1;
 CC EMBL; D11469; BAA02027.1;
 CC PIR; JQ2240; JQ2240.
 CC PIR; S46548; S46548.
 CC HSP; P02990; LEFU.
 CC InterPro; IPR004541; EF-Tu.
 CC InterPro; IPR004160; EFTU_Cterm.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR000795; EF-GTPbind.
 CC Pfam; PF00009; GTP_EFTU_D3; 1.
 CC Pfam; PF03143; GTP_EFTU_D3; 1.
 CC Pfam; PF03144; GTP_EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFC.
 CC TIGRFAMS; TIGR00231; small_GTP; 1.
 CC TIGRFAMS; TIGR00485; EF-Tu; 1.
 CC PROSITE; PS00301; EFATOR_GTP; 1.
 CC Elongation factor; Protein biosynthesis; GTP-binding;
 CC Transit peptide; Chloroplast.
 CC TRANSIT
 CC CHAIN 1 69 CHLOROPLAST (POTENTIAL).
 CC FT NP_BIND 70 478 ELONGATION FACTOR TU.
 CC FT NP_BIND 88 95 GTP (BY SIMILARITY).
 CC FT NP_BIND 150 154 GTP (BY SIMILARITY).
 CC FT NP_BIND 205 208 GTP (BY SIMILARITY).
 CC SQ SEQUENCE 478 AA; 51956 MW; 9811609E5FAC001 CRC64;
 CC
 CC Query Match 80.0%; Score 72; DB 1; Length 478;
 CC Best Local Similarity 78.9%; Pred. No. 4.3e-05;
 CC Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 1 RKGFEKTKPGVNTGTXV 19
 CC ||||| || ||||| |
 CC Db 72 RKGFEKTKPGVNTGTXV 90
 CC
 CC RESULT 5
 CC EFTL_SOYBN STANDARD; PRT; 479 AA.
 CC ID EFTL_SOYBN
 CC AC Q43467;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Elongation factor Tu, chloroplast precursor (EF-Tu).
 CC GN TUFA.
 CC OS Glycine max (Soybean).
 CC OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC OX euroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.

Plant Sci. 117:83-92(1996).
 -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
 AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
 BIOSYNTHESIS.
 -!- SUBCELLULAR LOCATION: Chloroplast.
 -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 EF-TU/EF-1A SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/entities>).

```

RC STRAIN=cv. Maple Arrow; TISSUE=Leaf;
RA Bony C., Stutz E.;
RT "Soymebean (Glycine max L.) nuclear DNA contains four tuf genes coding
RL for the chloroplast specific translation elongation factor EF-Tu.";
RL Chimia 47:247-249(1995)
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X66062; CAA46864.1; -.
CC HSSP; P07157; IAIP.
CC InterPro; IPR004541; EF-Tu.
CC InterPro; IPR004160; EFTU_Cterm.
CC InterPro; IPR004161; EFTU_D2.
CC InterPro; IPR000795; EF-GTPbind.
CC InterPro; IPR005225; Small-GTP.
CC InterPro; IPR00009; GTP-EFTU; 1.
CC Pfam; PF03143; GTP-EFTU_D3; 1.
CC Pfam; PF03144; GTP-EFTU_D2; 1.
CC PRINTS; PR00315; ELONGATNEFT.
CC TRIGRAMS; TIGR00231; small-GTP; 1.
CC TRIGRAMS; TIGR00485; EF-Tu; 1.
CC TRIGRAMS; TIGR00301; EFACOR-GTP; 1.
CC PROSITE; PS00301; EFACOR protein biosynthesis; GTP-binding;
CC Elongation factor, protein biosynthesis.
CC KW Elongation factor, protein biosynthesis.
CC KW Transit peptide; Chloroplast; Multigene family.
CC TRANSIT 1 71 CHLOROPLAST (POTENTIAL).
CC CHAIN 72 479 ELONGATION FACTOR TU.
CC FT NP_BIND 89 96 GTP (BY SIMILARITY).
CC FT NP_BIND 151 155 GTP (BY SIMILARITY).
CC FT NP_BIND 206 209 GTP (BY SIMILARITY).
CC FT CARBOHYD 333 333 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT SEQUENCE 479 AA; 52095 MW; 43C97B99EACFA971 CRC64;
CC SQ

```

[illegible]

the Swiss Institute of Bioinformatics. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL Outstation at the European Molecular Biology Laboratory. There are no restrictions on the use of this database. This statement as long as its contents are not used for commercial purposes requires a license agreement. Usage by an individual user requires a license agreement. Please send an email to license@ebi-sib.ch for commercial use. (See <http://www.isb.sib.ch/announcements>)

atch 77.8%; Score 70; DB 1; Length 46;
al Similarity 68.4%; Pred. No. 8.2e-06;
13; Conservative 3; Mismat.

DOSI
STANDARD:

ta sinensis (Marine centric diatom).
 a; stramenopiles; Bacillariophyta).
 iophycidae; Eupodiscales; Coscinodiscus
 ID-2839.

FEILER U.; *Biol. Rep.* 13:336-342(1995).
 TION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
 ACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROMOTING
 WITHSTHS.
 CELLULAR LOCATION: CHLOROPLAST, CYTOSOL, PLASTID, PLASTID
 MEMBRANE, PLASTID THYLAKOID, PLASTID THYLAKOID MEMBRANE, PLASTID

-----produced through a collaboration between the Center for Bioinformatics and the EMBL outstation at the Max Planck Institute for Biophysical Chemistry. There are no restrictions on the use of this statement as long as its content is in no way distorted. This document is made available under a Creative Commons license agreement (See <http://creativecommons.org/licenses/by/4.0/>)

0000795; EF-GTPbind.
005225; Small-GTP.
; GTP-EFTu; 1.



CC CC CC CC CC CC DR DR DR

Query
Best
Matched

RESULT 8
EFTU_ODOS
ID .EFTU
AC P494
DN

OG
OC
OC
OX
RN

CC
CC
CC
CC
CC
CC

CC use by
CC modified
CC entities
CC or send an
CC -----
DR FMO

DR
DR
DR

1

1



[illegible]

RESULT 10

RESULT 10		STANDARD:	PRT;	394 AA.
EFTU_RICPR				
EFTU_RICRP				
ID P4865:				
AC AC	01-FEB-1996	(Rel. 33, Created)		
DT DT	01-MAY-2000	(Rel. 39, Last sequence update)		
DT DT	30-MAY-2000	(Last annotation update)		
DT DT	16-OCT-2001	(Rel. 40, Last annotation factor Tu (EF-Tu).		
DE DE	Elongation factor Tu	(EF-Tu).		
GN GN	TUF OR RP661.			
OS OS	Rickettsia prowazekii:alpha subdivision; Rickettsiales:			
OC OC	Bacteria; Proteobacteri: Rickettsia.			
OC OC	Rickettsiaceae; Rickettsiaceae; Rickettsia.			
OX OX	NCBI_TaxID=782:			
[1]				
RN RN	SEQUENCE FROM N.A.			
RP RP	STRAIN=MADRID E;			
RC RC	MEDLINE=J7047977; PubMed=8892818;			
RX RX	Syvanen A., Amiri H., Jamal A., Andersson S.G.E., Kurland C.G.;			
RA RA	"A chimeric disposition of the elongation factor genes in Rickettsia			
PT PT	"A chimeric disposition of the elongation factor genes in Rickettsia			

prokazekit. 178:6192-0199A(1999).

J. Bacteriol.

[2] SEQUENCE FROM N.A.
STRAIN-Madrid E;
MEDLINE=039499; PubMed=9823693;
X MEDLINE=S.G.E., Zomorodipour A., Andersson J.O., Naeslund A.K.,
K Andersson S.G.E., Zomorodipour A., Podowski R.M., and the origin of
Andersson T., Alsmark U.C.M., Kurland C.G.;
A Sicheritz-Ponten T., Winkler H.H., Eriksson Ane sequence of Rickettsia prowazekii and the origin of
Rickettsia prowazekii.
A "The genome and sequence of Rickettsia prowazekii and the origin of
mitochondria".
NT Nature 396:133-140(1998). THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
TRL NATURE 396:133-140(1998). THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
-I- FUNCAACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC AMINOACID SYNTHESIS.
CC -I- SUBUNIT. MONOMER [BY SIMILARITY].
CC -I- SURCELLULAR LOCATION: Cytoplasmic.
CC -I- SURCELLULAR LOCATION: TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC -I- SIMILARITY: BELONGS TO THE EF-TU/EF-1A SUBFAMILY.
CC EF-TU/EF-1A SUBFAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute long as its content is in no way
CC derived from the EMBL outstation. Usage by non-profit institutions for commercial
CC use without prior agreement is not permitted. See http://www.isb-sib.ch/announce/
CC use policy and this statement agreement.)

[illegible]

AMINOACYL-TRNA
-1-
-1- GINOSYNTHESIS.
-1- SUBUNIT: MONOMER (BY SIMILARITY).
-1- SUBCELLULAR LOCATION: TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
-1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
-1- EF-TU/EF-1A SUBFAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is not removed, modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement> or send an email to license@sib-sib.ch).
EMBL: Z34275; CAA84029.1;
EMBL: AE00251; RAF30935.1;
HSSP: P02990; 1FTU. EF-TU
InterPro: IPR004160; EFTU_D2
InterPro: IPR004161; EFTU_D1
InterPro: IPR000795; EF-GTPbind
InterPro: IPR000795; EF-GTPbind


```
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PR00315; ELONGATNFT.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
DR Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 CTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 394 AA; 42902 MW; 80A887B6C59883E0 CRC64;

Query Match 76.7%; Score 69; DB 1; Length 394;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
DB 3 KAKFERTKPHVNIQTIGHV 21

RESULT 12
EFTU_BACST
ID EFTU_BACST STANDARD; PRT; 395 AA.
AC O50306;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 2184;
RX MEDLINE=98443240; PubMed=9769211;
RA Krasny L., Mesters J.R., Tieleman L.N., Kraal B., Fucik V.,
RA Hilgenfeld R., Jonak J.;
RT "Structure and expression of elongation factor Tu from Bacillus
RT stearothermophilus.";
RL J. Mol. Biol. 283:371-381(1998).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ000260; CAA03976.1;
CC HSP: P02990; IETU.
CC InterPro: IPR004541; EF-Tu.
CC InterPro: IPR004160; EFTU_Cterm.
CC InterPro: IPR000795; EF_GTPbind.
CC InterPro: IPR005225; Small_GTP.
CC Pfam: PF00009; GTP_EFTU; 1.
CC Pfam: PF03143; GTP_EFTU_D3; 1.
CC Pfam: PF03144; GTP_EFTU_D2; 1.
CC PRINTS: PR00315; ELONGATNFT.
CC TIGRFAMS: TIGR00231; small_GTP; 1.
```

```
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
DR Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 395 AA; 43290 MW; 71B08165E7FB42C5 CRC64;

Query Match 76.7%; Score 69; DB 1; Length 395;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
DB 3 KAKFERTKPHVNIQTIGHV 21

RESULT 13
EFTU_CORGL
ID EFTU_CORGL STANDARD; PRT; 396 AA.
AC P42439;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR CGL0497.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13059 / AS019;
RX MEDLINE=94368062; PubMed=8085791;
RA Ludwig W., Neumaier J., Klugbauer N., Brockmann E., Rollier C.,
RA Klugbauer S., Reetz K., Schachtner I., Ludwigsen A.,
RA Bachleitner M., Fischer U., Schleifer K.H.;
RT "Phylogenetic relationships of Bacteria based on comparative sequence
RT analysis of elongation factor Tu and ATP-synthase beta-subunit
RT genes.";
RL Antonie Van Leeuwenhoek 64:285-305(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX Nakagawa S.;
RA "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X77034; CAA54323.1;
CC EMBL: AP005275; BAB97890.1;
CC HSP: P02990; IETU.
CC InterPro: IPR004541; EF-Tu.
CC InterPro: IPR004160; EFTU_Cterm.
CC InterPro: IPR004161; EFTU_D2.
CC InterPro: IPR000795; EF_GTPbind.
CC InterPro: IPR005225; Small_GTP.
CC Pfam: PF00009; GTP_EFTU; 1.
```

```

DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNCT.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR TIGRFAMS; TIGR00485; EF-Tu; 1.
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
SQ SEQUENCE 396 AA; 43852 MW; F46333EC113AB5FD CRC64;

Query Match 76.7%; Score 69; DB 1; Length 396;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
DB 3 KAKFERTKPHVNIQTIGHV 21

RESULT 14
EFTU_MYCLE
ID EFTU_MYCLE STANDARD; PRT; 396 AA.
AC P30768;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR MLI877.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.; PubMed=8341612;
RA MEDLINE=93341950; PubMed=8341612;
RA Silbak F., Bercovier H.;
RT "Nucleotide sequence of Mycobacterium leprae elongation factor
(EF-Tu) gene.";
RL Nucleic Acids Res. 21:3327-3327(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93188701; PubMed=8446028;
RA Honore N.T., Bergh S., Chanteau S., Doucet-Populaire F.,
RA Eigmeier K., Garnier T., Georges C., Launois P., Limpalaboon T.,
RA Newton S., Nsang K., del Portillo P., Ramesh G.R., Reddi P.,
RA Ridei F.R., Sittisombut N., Wu-Hunter S., Cole S.T.;
RT "Nucleotide sequence of the first cosmid from the Mycobacterium
leprae genome project: structure and function of the Rif-Str
regions.";
RL Mol. Microbiol. 7:207-214(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94375410; PubMed=8089081;
RA Dhandayathapani S., Banu J.M., Kashiwabara Y.;
RT "Cloning and sequence determination of the gene coding for the
elongation factor Tu of Mycobacterium leprae.";
RL J. Biochem. 115:664-669(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Mungall K., Basham D., Brown D., Churcher C., Harris D.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd L., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy S., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";

```

```

RL Nature 409:1007-1011(2001).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L13276; AAA71969.1; -
DR EMBL; Z14314; CAA78674.1; -
DR EMBL; D13869; BAA02982.2; -
DR EMBL; AL583923; CAC30831.1; -
DR PIR; S34954; S34954.
DR HSP; P02990; LETU.
DR Leproma; MLI877; -
DR InterPro; IPR004541; EF-Tu.
DR InterPro; IPR004160; EFTU_Cterm.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR TIGRFAMS; TIGR00485; EF-Tu; 1.
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
FT CONFLICT 129 152 GPPYLLVALKSDAVDELELLELV -> VYLTSWHLTSPT
FT CONFLICT 281 281 PWTRNYSLS (IN REF. 2).
FT CONFLICT 349 349 L -> V (IN REF. 2).
FT CONFLICT 384 384 G -> A (IN REF. 2).
FT CONFLICT 384 384 R -> P (IN REF. 3).
SQ SEQUENCE 396 AA; 43667 MW; D9CB88343C642778 CRC64;

Query Match 76.7%; Score 69; DB 1; Length 396;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIQTIXV 19
DB 3 KAKFERTKPHVNIQTIGHV 21

RESULT 15
EFTU_STRCJ
ID EFTU_STRCJ STANDARD; PRT; 397 AA.
AC P95724;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF.
OS Streptomyces cinnamonensis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=53446;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUE89;
RA Cappellano C., Monti F., Sosio M., Donadio S., Sarubbi E.;

```

Fri May 2 08:51:22 2003

Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.

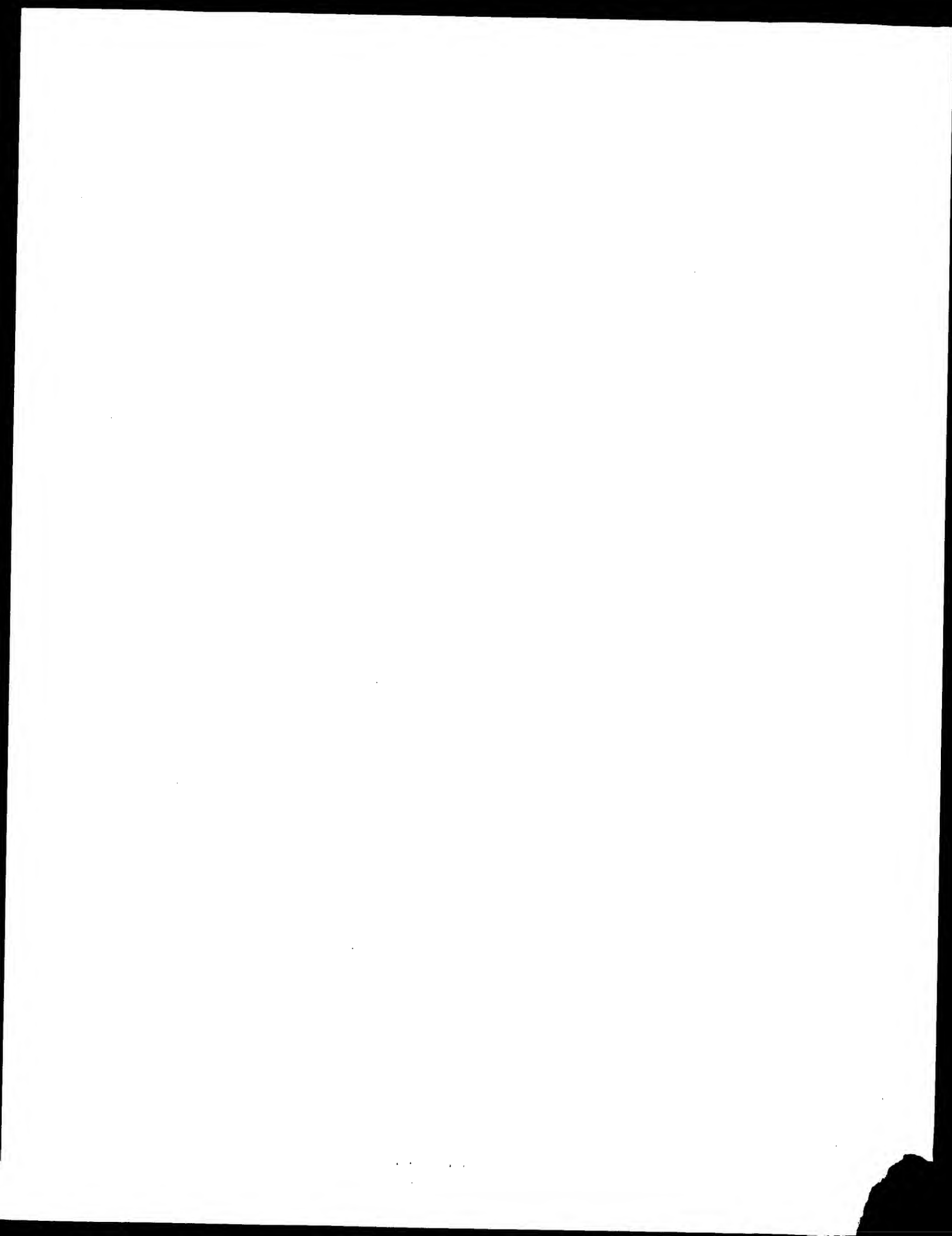
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X98831; CAA67349.1; -
CC HSSP; P02990; LETU.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF01143; GTP_EFTU_D3; 1.
DR Pfam; PF01144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFACT.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR TIGRFAMS; TIGR00485; EF-Tu; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
ET NP_BIND 138 141 GTP (BY SIMILARITY).
SQ SEQUENCE 397 AA; 43891 MW; EEB8235730D8BBB7 CRC64;

Query Match 76.7%; Score 69; DB 1; Length 397;
Best Local Similarity 73.7%; Pred. No. 0.00012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RCKFERTKPGVNICTIXAV 19
: ||||| ||||| |
Db 3 KAKFERTKPHVNICTIGHV 21

Search completed: April 29, 2003, 09:40:54
Job time : 26 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:36:44 ; Search time 29 Seconds
(without alignments)
134.996 Million cell updates/sec

Title: US-09-810-764A-5

Perfect score: 90

Sequence: 1 RGFERTKPGVNGTIXV 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriophage.*

17: sp_archaeal.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	86.7	467	10 Q9SEF8	Q9sef8 oryza sativ
2	78	86.7	467	10 Q8W2C3	Q8w2c3 oryza sativ
3	75	83.3	391	2 Q9F9S8	Q9f9s8 edta-degrad
4	72	80.0	409	16 Q8YP63	Q8yp63 anabaena sp
5	72	80.0	457	10 Q40450	Q40450 nicotiana s
6	72	80.0	474	10 Q9AXU2	Q9axu2 pelargonium
7	72	80.0	485	10 Q43364	Q43364 nicotiana s
8	69	76.7	391	16 Q8YHP2	Q8yhp2 bruceella me
9	69	76.7	394	2 Q933R9	Q933r9 ureaplasma
10	69	76.7	394	16 Q92GW4	Q92gw4 rickettsia
11	69	76.7	397	2 Q9RG55	Q9rg55 streptomyces
12	69	76.7	406	16 Q8YHQ4	Q8yhq4 bruceella me
13	68	75.6	397	2 Q9RG53	Q9rg53 streptomyces
14	68	75.6	397	2 Q93T39	Q93t39 streptomyces
15	67	74.4	397	16 Q97EH5	Q97eh5 clostridium
16	67	74.4	410	8 Q9TLV8	Q9tlv8 cyanidium c

17	66	73.3	394	16 Q9KV37	Q9kv37 vibrio chol
18	66	73.3	394	16 Q9KU26	Q9kuz6 vibrio chol
19	66	73.3	394	16 Q8ZJB2	Q8zjb2 versinia pe
20	66	73.3	394	16 Q8ZAN8	Q8zan8 versinia pe
21	66	73.3	394	16 Q8XED3	Q8xed3 escherichia
22	66	73.3	394	16 Q8X4S9	Q8x4s9 escherichia
23	66	73.3	396	16 Q8XGZ0	Q8xgz0 raistonia s
24	66	73.3	400	16 Q8R7V2	Q8r7v2 thermoanaer
25	66	73.3	400	16 Q8R7T8	Q8r7t8 thermoanaer
26	66	73.3	409	8 Q9TJQ8	Q9tjq8 prototheca
27	66	73.3	409	8 Q9TJQ8	Q9tjq8 prototheca
28	64	71.1	391	16 Q8XFF8	Q8xfl8 clostridium
29	64	71.1	391	16 Q8UE16	Q8uel6 agrobacteri
30	63	70.0	393	2 Q925Y6	Q925y6 rhizobium m
31	62	68.9	394	16 Q9RAJ9	Q9raj9 moraxella s
32	62	68.9	394	16 Q9K117	Q9kil7 neisseria m
33	62	68.9	394	16 Q9JR15	Q9jri5 neisseria m
34	62	68.9	396	16 Q99QW0	Q99qm0 caulobacter
35	60	66.7	394	3 Q8TGG6	Q8tgg6 aspergillus
36	60	66.7	395	16 Q8R603	Q8rg66 fusobacteri
37	60	66.7	395	16 Q927I6	Q927i6 listeria in
38	59	65.6	437	3 Q9C2P1	Q9c2p1 listeria mo
39	57	63.3	398	16 Q97PV3	Q97pv3 streptococc
40	56	62.2	300	5 Q95TV3	Q95tv3 drosophila
41	56	62.2	401	2 Q9XD38	Q9xd38 leptospira
42	56	62.2	489	5 Q9V6Q7	Q9v6q7 drosophila
43	55	61.1	453	10 Q8W2C4	Q8w2c4 oryza sativ
44	53	58.9	358	2 Q9X6G1	Q9x6g1 coxiella bu
45	52	57.8	395	2 Q9R420	Q9r420 porphyromon

ALIGNMENTS

RESULT 1

ID	Q9SEF8	PRELIMINARY;	PRT;	467 AA.
AC	Q9SEF8;			
DT	01-MAY-2000	(TrEMBLrel. 13, Created)		
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)		
DE	Chloroplast translational elongation factor Tu.			
GN	TUFA.			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzaceae; Oryza.			
OX	NCBI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV. JAPONICA NIPPONBARE;			
RX	MEDLINE-20062172; PubMed=10597036;			
RA	Lee J.H., Lee J.W., Chung Y.Y., Paek K.H., Shin J.S., Yun C.H.,			
RA	Kim J.K.;			
RT	"Cloning and characterization of the chloroplast elongation factor EF-			
RL	Tu cDNA of Oryza sativa L.,"			
DR	Mol. Cells 9:484-490(1999).			
DR	EMBL; AF145053; AAF15312.1; .			
DR	HSSP; P07157; IAIIP.			
DR	InterPro; IPR004541; EF-Tu.			
DR	InterPro; IPR004160; EFTU_Cterm.			
DR	InterPro; IPR004161; EFTU_D2.			
DR	InterPro; IPR000795; EF_GTPbind.			
DR	InterPro; IPR005225; Small_GTP.			
DR	Pfam; PF00009; GTP_EFTU_1.			
DR	Pfam; PF03144; GTP_EFTU_D2; 1.			
DR	Pfam; PF03143; GTP_EFTU_D3; 1.			
DR	PRINTS; PR00315; ELONGATNFCT.			
DR	TIGRFAMS; TIGR00485; EF-Tu; 1.			
DR	TIGRFAMS; TIGR00231; small_GTP; 1.			
DR	PROSITE; PS00301; EFATOR_GTP; 1.			
KW	Elongation factor; GTP-binding; Protein biosynthesis.			
KW	Elongation factor; GTP-binding; Protein biosynthesis.			
SEQUENCE	467 AA; 50354 MW; 79A119279A7F1847 CRC64;			

```

Query Match      86.7%; Score 78; DB 10; Length 467;
Best Local Similarity 84.2%; Pred. No. 2.3e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFKRTKPGVNICTIXXV 19
DB 61 RGKFKRTKPGVNICTIGHV 79

RESULT 2
Q8W2C3 PRELIMINARY; PRT; 467 AA.
AC Q8W2C3: 2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Translational elongation factor Tu.
GN TUFA.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-H., Kim J.-K.;
RT "Chloroplast translational elongation factor Tu gene of rice.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327413; AAL37431.1; -.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF_GTPbind.
DR Pfam; PF000009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMS; TIGR00485; EF-Tu; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor.
SQ SEQUENCE 467 AA; 50413 MW; BA6D9FEABDF02CC4 CRC64;

Query Match      86.7%; Score 78; DB 10; Length 467;
Best Local Similarity 84.2%; Pred. No. 2.3e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFKRTKPGVNICTIXXV 19
DB 61 RGKFKRTKPGVNICTIGHV 79

RESULT 3
Q9F9S8 PRELIMINARY; PRT; 391 AA.
AC Q9F9S8: 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Eftu.
GN EFTU.
OS EDTA-degrading bacterium BNCL.
OC Bacteria; Proteobacteria; alpha subdivision.
OX NCBI_TaxID=85561;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BNCL;
RX MEDLINE=21091958; PubMed=11157232;
RA Bohuslavsk J., Payne J.W., Liu Y., Bolton H. Jr., Xun L.;
RT "Cloning, Sequencing, and Characterization of a Gene Cluster Involved
in EDTA Degradation from the Bacterium BNCL.";
RL Appl. Environ. Microbiol. 67:688-695(2001).
RN [2]

```

```

SEQUENCE FROM N.A.
RC STRAIN=BNCL;
RX MEDLINE=21091959; PubMed=11157233;
RA Liu Y., Louie T.M., Payne J., Bohuslavsk J., Bolton H. Jr., Xun L.;
RT "Identification, purification, and characterization of iminodiacetate
oxidase from the EDTA-degrading bacterium BNCL.";
RL Appl. Environ. Microbiol. 67:696-701(2001).
DR EMBL; AF176664; RAG09263.1; -.
DR HSSP; P07157; IALP.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam; PF000009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMS; TIGR00485; EF-Tu; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW GTP-binding.
SQ SEQUENCE 391 AA; 42794 MW; E6B41737CDD77AA6 CRC64;

Query Match      83.3%; Score 75; DB 2; Length 391;
Best Local Similarity 78.9%; Pred. No. 6.5e-05;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFKRTKPGVNICTIXXV 19
DB 3 KGKFKRTKPGVNICTIGHV 21

RESULT 4
Q8YP63 PRELIMINARY; PRT; 409 AA.
AC Q8YP63: 2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Translation elongation factor EF-Tu.
GN TUFA OR ALL4337.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AF003596; BAB76036.1; -.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam; PF000009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMS; TIGR00485; EF-Tu; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; Complete proteome.
SQ SEQUENCE 409 AA; 44811 MW; 084708FBA8ACD9D3 CRC64;

Query Match      80.0%; Score 72; DB 16; Length 409;

```

Best Local Similarity 78.9%; Pred. No. 0.00023;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RCKFERTKPGVNIPTIXV 19
Db 3 RAKFERTKPHVNIPTIGHV 21

RESULT 5

Q04050 ID Q04050 PRELIMINARY; PRT; 457 AA.
AC Q04050
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chloroplast elongation factor TUA (EF-TUA) precursor (Fragment).
GN TUFA.
OS Nicotiana sylvestris (Wood tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4096;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA MEDLINE=93363910; PubMed=8358028;
RX Murayama Y., Matsubayashi T., Sugita M., Sugliura M.;
RT "Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco: the existence of two chloroplast elongation factor Tu species.";
RL Plant Mol. Biol. 22:767-774(1993).
DR EMBL; D11375; BAA01974.1; -.
DR HSSP; P02990; LEFU.
DR InterPro; IPR004541; EF-Tu.
DR InterPro; IPR004160; EFTU_Cterm.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR00795; EF_GTPhind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR TIGRFAMs; TIGR00485; EF-Tu; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis.
FT TRANSIT 1 70
FT CHAIN 71 >457
FT NON_TER 457 457
FT SEQUENCE 457 AA; 49732 MW; 8A30C50DC137F755 CRC64;

Query Match 80.0%; Score 72; DB 10; Length 457;
Best Local Similarity 78.9%; Pred. No. 0.00027;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RCKFERTKPGVNIPTIXV 19
Db 72 RCKFERTKPHVNIPTIGHV 90

RESULT 6

Q9AXU2 ID Q9AXU2 PRELIMINARY; PRT; 474 AA.
AC Q9AXU2
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chloroplast translational elongation factor Tu.
GN TUFA.
OS Pelargonium graveolens (rose geranium).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Geraniales; Geraniaceae; Pelargonium.

OX NCBI_TaxID=73200;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20550852; PubMed=11101151;
RA Kang C.J., Lee M.G., Cho Y.S., Lee J.W., Kyung Y.J., Shin J.S.,
RA Kim E.S., Kim J.K.;
RT "Characterization of geranium (Pelargonium graveolens) chloroplast EF-Tu cDNA.";
RT Tu cDNA.";
RL Mol. Cells 10:579-583(2000).
DR EMBL; AF234537; AAK08141.1; -.
DR HSSP; P02990; LEFU.
DR InterPro; IPR004541; EF-Tu.
DR InterPro; IPR004160; EFTU_Cterm.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR00795; EF_GTPhind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00485; EF-Tu; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis.
SQ SEQUENCE 474 AA; 51315 MW; A321742468553B65 CRC64;

Query Match 80.0%; Score 72; DB 10; Length 474;
Best Local Similarity 78.9%; Pred. No. 0.00028;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RCKFERTKPGVNIPTIXV 19
Db 68 RCKFERTKPHVNIPTIGHV 86

RESULT 7

Q43364 ID Q43364 PRELIMINARY; PRT; 485 AA.
AC Q43364
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chloroplast elongation factor TUB(EF-TUB) precursor.
GN TUFA OR TUBF.
OS Nicotiana sylvestris (Wood tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94373864; PubMed=8087886;
RA Sugita M., Murayama Y., Sugliura M.;
RT "Structure and differential expression of two distinct genes encoding the chloroplast elongation factor Tu in tobacco.";
RL Curr. Genet. 25:164-168(1994).
RN [2]
RP SEQUENCE OF 61-485 FROM N.A.
RC TISSUE=YOUNG LEAF;
RX MEDLINE=93363910; PubMed=8358028;
RA Murayama Y., Matsubayashi T., Sugita M., Sugliura M.;
RT "Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco: the existence of two chloroplast elongation factor Tu species.";
RL Plant Mol. Biol. 22:767-774(1993).
DR EMBL; D11470; BAA02028.1; -.
DR EMBL; D11376; BAA01975.1; -.
DR HSSP; P02990; LEFU.
DR InterPro; IPR004541; EF-Tu.
DR InterPro; IPR004160; EFTU_Cterm.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR00795; EF_GTPhind.
DR InterPro; IPR005225; Small_GTP.

Fri May 2 08:51:23 2003

DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR PRINTS; PR00315; ELONGATNFACT.
 DR TIGRFAMS; TIGR00485; EF-Tu; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 DR Elongation factor; GTP-binding; Protein biosynthesis; Transit peptide.
 KW TRANSIT 1 77 POTENTIAL.
 FT CHAIN 78 485 CHLOROPLAST ELONGATION FACTOR TUB
 FT SEQUENCE 485 AA; 52688 MW; FE88EF5D48B92D29 CRC64;
 SQ

Query Match 80.0%; Score 72; DB 10; Length 485;
 Best Local Similarity 78.9%; Pred. No. 0.00028;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKRTKPGVNIPTIXV 19
 DQ 79 RGKFKRKPHVNIPTIGHV 97

RESULT 8
 QYHP2
 ID Q8YHP2 PRELIMINARY; PRT; 391 AA.
 AC Q8YHP2;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Protein translation elongation factor Tu (EF-Tu).
 GN BME10755.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujar C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kyrpides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis";
 Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 RL EMBL; AE009518; AAL51936.1; -
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR TIGRFAMS; TIGR00485; EF-Tu; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 42604 MW; ED4CDF37183A900E CRC64;

Query Match 76.7%; Score 69; DB 16; Length 391;
 Best Local Similarity 73.7%; Pred. No. 0.00076;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKRTKPGVNIPTIXV 19
 DQ 3 KSKFKRKPHVNIPTIGHV 21

RESULT 9

Q933R9
 ID Q933R9 PRELIMINARY; PRT; 394 AA.
 AC Q933R9;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Elongation factor Tu.
 GN TUF.
 OS Ureaplasma urealyticum (Ureaplasma urealyticum biotype 2).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=21130;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-VARIOUS STRAINS;
 RA Kong F., Gilbert G.L.;
 RT "Genomic based mapping and sequencing of U. parvum and U.
 urealyticum";
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF270767; AAK58630.1; -
 DR EMBL; AF270758; AAK58621.1; -
 DR EMBL; AF270759; AAK58622.1; -
 DR EMBL; AF270760; AAK58623.1; -
 DR EMBL; AF270761; AAK58624.1; -
 DR EMBL; AF270762; AAK58625.1; -
 DR EMBL; AF270763; AAK58626.1; -
 DR EMBL; AF270764; AAK58627.1; -
 DR EMBL; AF270765; AAK58628.1; -
 DR EMBL; AF270766; AAK58629.1; -
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR TIGRFAMS; TIGR00485; EF-Tu; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; UNKNOWN_1.
 KW GTP-binding.
 SQ SEQUENCE 394 AA; 42918 MW; 4CB4B2D776A5B145 CRC64;

Query Match 76.7%; Score 69; DB 2; Length 394;
 Best Local Similarity 73.7%; Pred. No. 0.00077;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKRTKPGVNIPTIXV 19
 DQ 3 KAKFKRKPHVNIPTIGHV 21

RESULT 10
 Q92GW4
 ID Q92GW4 PRELIMINARY; PRT; 394 AA.
 AC Q92GW4;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Elongation factor EF-Tu.
 GN TUF OR RC1008.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsia.
 OX NCBI_TaxID=781;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=WALISH 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";

Science 293:2093-2098(2001).
 EMBL; AE008652; AAL03546.1; EF-Tu.
 InterPro; IPR004541; EF-Tu.
 InterPro; IPR004160; EFTU_Cterm.
 InterPro; IPR004161; EFTU_D2.
 InterPro; IPR000795; EF_GTPbind.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00009; GTP_EFTU; 1.
 Pfam; PF03144; GTP_EFTU_D2; 1.
 Pfam; PF03143; GTP_EFTU_D3; 1.
 TIGRFAMs; TIGR00485; EF-Tu; 1.
 TIGRFAMs; TIGR00231; small_GTP; 1.
 PROSITE; PS00301; EFATOR_GTP; UNKNOWN_1.
 Complete proteome.
 NCBI_TaxID=29459;
 SEQUENCE 394 AA; 42868 MW; 1FBB465785530C63 CRC64;

Query Match 76.7%; Score 69; DB 16; Length 394;
 Best Local Similarity 73.7%; Pred. No. 0.00077;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIPTIXXV 19
 : ||||| |||||
 Db 3 KAKFERTKPHVNIPTIGHV 21

RESULT 11

Q9RG55 PRELIMINARY; PRT; 397 AA.
 AC Q9RG55;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Elongation factor Tu1.
 GN TU1.

OS Streptomyces mobaraensis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=35621;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC29032;
 RA Olsthoorn-Tielemann L.N., Claaij N., Hilgenfeld R., Kraal B.;
 RT "Elongation factor Tu1 from the pulvomycin producer Streptomycetes mobaraensis is resistant to both pulvomycin and kirromycin.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF153617; AAF22606.1; -;
 DR HSSP; P02990; 1ETU.
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00485; EF-Tu; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 KW GTP-binding.
 SQ SEQUENCE 397 AA; 43811 MW; DA439151BFE6BC2E CRC64;

Query Match 76.7%; Score 69; DB 2; Length 397;
 Best Local Similarity 73.7%; Pred. No. 0.00078;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIPTIXXV 19
 : ||||| |||||
 Db 3 KAKFERTKPHVNIPTIGHV 21

RESULT 12

Q8YHQ4

Q8YHQ4 PRELIMINARY; PRT; 406 AA.
 Q8YHQ4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Protein translation elongation factor Tu (EF-Tu).
 GN EME10742.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Muej C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kypides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL; AE009516; AAL51923.1; -;
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00485; EF-Tu; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 KW Complete proteome.
 SQ SEQUENCE 406 AA; 44340 MW; 9523841BFF2C08DF CRC64;

Query Match 76.7%; Score 69; DB 16; Length 406;
 Best Local Similarity 73.7%; Pred. No. 0.0008;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNIPTIXXV 19
 : ||||| |||||
 Db 18 KSKFERTKPHVNIPTIGHV 36

RESULT 13

Q9RG53 PRELIMINARY; PRT; 397 AA.
 ID Q9RG53;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Elongation factor Tu1.
 GN TU1.

OS Streptococcus pneumoniae (Streptococcus pneumoniae).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=55404;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TU1063;
 RA Olsthoorn-Tielemann L.N., Claaij N., Hilgenfeld R., Kraal B.;
 RT "Elongation factor Tu1 from the pulvomycin producer Streptomycetes mobaraensis is resistant to both pulvomycin and kirromycin.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF153618; AAF22608.1; -;
 DR HSSP; P02990; 1ETU.
 DR InterPro; IPR004541; EF-Tu.
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR004161; EFTU_D2.

DR InterPro: IPR000795; EF-GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU_1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
DR GTP-binding_397 AA; 43682 MW; 5F3A81F3FC42914F CRC64;
KW SEQUENCE 397 AA; 43682 MW; 5F3A81F3FC42914F CRC64;
SQ
Query Match 75.6%; Score 68; DB 2; Length 397;
Best Local Similarity 68.4%; Pred. No. 0.0012;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 RGFERTKPGVNIQTIXV 19
DB : ||||| ||||| :
3 KAKFERTKPHVNIQTIGHI 21
[1]
RESULT 14
Q3T39 PRELIMINARY; PRT; 397 AA.
AC Q3T39;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Elongation factor Tu.
GN TUF1.
OS Streptomyces aureofaciens.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=84/25;
RA Kormanec J., Nguyen L.D., Novotna J., Knirschova R., Weiser J.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF368284; AKS54131.1; -.
DR InterPro: IPR004541; EFTU_1.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF-GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
DR GTP-binding_397 AA; 43683 MW; BF107EA8036FEF82 CRC64;
KW SEQUENCE 397 AA; 43683 MW; BF107EA8036FEF82 CRC64;
SQ
Query Match 75.6%; Score 68; DB 2; Length 397;
Best Local Similarity 68.4%; Pred. No. 0.0012;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 RGFERTKPGVNIQTIXV 19
DB : ||||| ||||| :
3 KAKFERTKPHVNIQTIGHI 21
[1]
RESULT 15
Q97EH5 PRELIMINARY; PRT; 397 AA.
AC Q97EH5;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Elongation factor Tu (Ef-Tu).
GN CAC3136.
OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Oiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Koonin E.V., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007809; AKK81075.1; -.
DR InterPro: IPR004541; EFTU_Cterm.
DR InterPro: IPR004160; EFTU_D2.
DR InterPro: IPR004161; EFTU_D3.
DR InterPro: IPR000795; EF-GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 397 AA; 43452 MW; F031B2F700434452 CRC64;
Query Match 74.4%; Score 67; DB 16; Length 397;
Best Local Similarity 82.4%; Pred. No. 0.0018;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 KFERTKPGVNIQTIXV 19
DB : ||||| ||||| :
5 KFERTKPHVNIQTIGHV 21
[1]
Search completed: April 29, 2003, 09:41:32
Job time : 32 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:39:04 ; Search time 29 Seconds
(without alignments)
19.277 Million cell updates/sec

Title: US-09-810-764A-5

Perfect score: 90

Sequence: 1 RGFERTKPGVNIQTIXV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	75.6	409	4	US-09-140-466-3
2	67	74.4	408	4	US-09-140-466-4
3	66	73.3	394	4	US-09-140-466-5
4	55	61.1	409	4	US-09-140-466-5
5	48	53.3	394	4	US-09-218-197-2
6	48	53.3	414	4	US-09-134-001C-5377
7	41	45.6	861	1	US-08-764-100-24
8	40	44.4	406	1	US-07-973-431B-1
9	39	43.3	260	4	US-09-254-465A-23
10	39	43.3	263	4	US-09-254-465A-25
11	39	43.3	299	4	US-09-188-930-189
12	39	43.3	299	4	US-09-462-270-2
13	39	43.3	299	4	US-09-462-270-2
14	39	43.3	299	4	US-09-254-465A-1
15	39	43.3	739	4	US-08-444-818-148
16	39	43.3	2955	2	US-08-443-260-3
17	39	43.3	2955	3	US-08-442-805A-3
18	39	43.3	2955	3	US-08-443-900A-3
19	39	43.3	2955	4	US-08-444-818-124
20	39	43.3	2955	4	US-08-249-843-3
21	39	43.3	3011	2	US-08-444-818-138
22	39	43.3	3011	4	US-08-833-678A-6
23	39	43.3	3011	4	US-08-444-818-177
24	39	43.3	3011	4	US-08-529-169A-6
25	38	42.2	255	1	US-08-152-019A-36
26	38	42.2	410	4	US-09-140-466-2
27	38	42.2	461	4	US-09-134-001C-3316

```

28 38 42.2 1112 2 US-08-714-402-2 Sequence 2, Appli
29 38 42.2 1161 4 US-09-327-536-2 Sequence 2, Appli
30 37 41.1 330 2 US-08-997-080-30 Sequence 30, Appl
31 37 41.1 330 2 US-08-997-362-30 Sequence 30, Appl
32 37 41.1 330 3 US-08-873-970-30 Sequence 30, Appl
33 37 41.1 330 4 US-09-095-855-30 Sequence 30, Appl
34 37 41.1 330 4 US-08-705-347A-30 Sequence 30, Appl
35 37 41.1 330 4 US-09-324-542-30 Sequence 30, Appl
36 37 41.1 330 4 US-09-205-426-30 Sequence 30, Appl
37 37 41.1 330 4 US-09-200-643-30 Sequence 30, Appl
38 37 41.1 1239 2 US-08-937-931-2 Sequence 2, Appli
39 37 41.1 1239 4 US-09-285-502-2 Sequence 2, Appli
40 37 41.1 1239 4 US-09-709-126-2 Sequence 2, Appli
41 37 41.1 1239 4 US-09-871-385A-2 Sequence 2, Appli
42 36.5 40.6 765 4 US-08-444-818-70 Sequence 70, Appl
43 36 40.0 28 1 US-07-696-551B-1 Sequence 1, Appli
44 36 40.0 315 4 US-09-154-750A-91 Sequence 91, Appl
45 35 38.9 185 2 US-08-770-544-18 Sequence 18, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-140-466-3
; Sequence 3, Application US/09140466
; Patent No. 6268160
; GENERAL INFORMATION:
; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; FILE REFERENCE: N68837B GCW PJC DP
; CURRENT APPLICATION NUMBER: US/09/140,466
; CURRENT FILING DATE: 1998-08-26
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Anacystis nidulans
US-09-140-466-3

```

```

Query Match 75.6%; Score 68; DB 4; Length 409;
Best Local Similarity 73.7%; Pred. No. 0.0012;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 RGFERTKPGVNIQTIXV 19
DB 3 RAKFERTKPHANIGTICHV 21

```

```

RESULT 2
US-09-140-466-4
; Sequence 4, Application US/09140466
; Patent No. 6268160
; GENERAL INFORMATION:
; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; FILE REFERENCE: N68837B GCW PJC DP
; CURRENT APPLICATION NUMBER: US/09/140,466
; CURRENT FILING DATE: 1998-08-26
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 14

```


APPLICANT: van Grinsven J., Martinus Q.
APPLICANT: De Haan, Petrus T.
APPLICANT: Gielen L., Johannes J.
APPLICANT: Peters, Dirk
APPLICANT: Goldbach, Robert W.
TITLE OF INVENTION: Improvements in or Relating to Organic
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,100
FILING DATE: 06-DEC-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,064
FILING DATE:
APPLICATION NUMBER: US 08/032,235
FILING DATE: 17-MAR-1993
APPLICATION NUMBER: GB 9206016.9
FILING DATE: 19-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5773700ris, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER: 137-1061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 354-3592
TELEFAX: (415) 857-1125
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: protein
MOLECULE TYPE: protein
US-08-764-100-24

Query Match 45.6%; Score 41; DB 1; Length 861;
Best Local Similarity 61.5%; Pred. No. 74;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGFERTKPGVNI 13
|| || || || ||
DB 203 RGLSERASPLNI 215

RESULT 8
US-07-973-431B-1
Sequence 1, Application US/07973431B
Patent No. 5652144
GENERAL INFORMATION:
APPLICANT: Lu, Yinchun
TITLE OF INVENTION: YC1 Protein, Gene, And Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: David G. Conlin; Dike, Bronstein,
ADDRESSEE: Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/973,431B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Eisenstein, Ronald I
REGISTRATION NUMBER: 30628
REFERENCE/DOCKET NUMBER: 41968
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 406 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-07-973-431B-1

Query Match 44.4%; Score 40; DB 1; Length 406;
Best Local Similarity 57.1%; Pred. No. 50;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GRKFERTKPGVNIGT 15
||| ||| |||
DB 203 CKFIKTPGVNSAPT 216

RESULT 9
US-09-254-465A-23
Sequence 23, Application US/09254465A
Patent No. 6410708
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Fong, Sherman
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Napier, Mary A.
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William I.
TITLE OF INVENTION: COMPOUNDS, COMPOSITIONS AND METHODS FOR THE TREATMENT
OF DISEASES CHARACTERIZED BY A33-RELATED ANTIGENS
FILE REFERENCE: P1216R1(US)
CURRENT APPLICATION NUMBER: US/09/254,465A
CURRENT FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: PCT/US98/24855
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: US 60/066,364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: US 60/078,936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 30
SEQ ID NO 23
LENGTH: 260
TYPE: PRT
ORGANISM: Homo sapiens
US-09-254-465A-23

Query Match 43.3%; Score 39; DB 4; Length 260;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGFERTKPGVN 12
|| || || || ||

; Sequence 1, Application US/09254465A
; Patent No. 6410708
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Fong, Sherman
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Napier, Mary A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOUNDS, COMPOSITIONS AND METHODS FOR THE TREATMENT
; OF DISEASES CHARACTERIZED BY A33- RELATED ANTIGENS
; FILE REFERENCE: P1216R1(US)
; CURRENT APPLICATION NUMBER: US/09/254,465A
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: PCT/US98/24855
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: US 60/066,364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 60/078,936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: PCT/US98/19437
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 30
; SEQ ID NO 1
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-465A-1

Query Match 43.3%; Score 39; DB 4; Length 299;
Best Local Similarity 58.3%; Pred. No. 54;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFERTKPGVN 12
|||:|||||:
Db 263 RGHFDRKKGTS 274

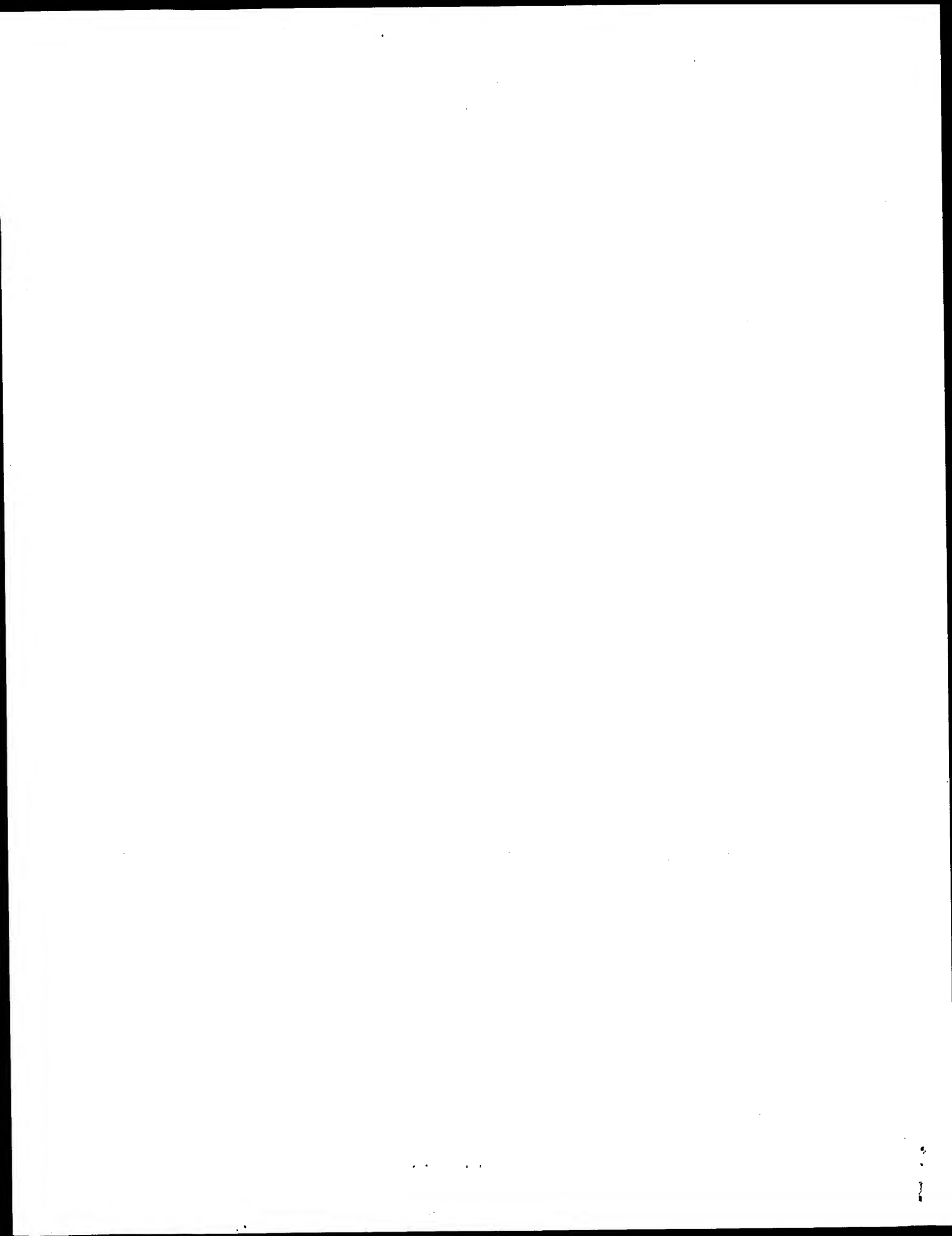
RESULT 15
US-08-444-818-148
; Sequence 148, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-444-818-148

Query Match 43.3%; Score 39; DB 4; Length 739;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RGKFERTKPGVN 12
|||:|||||:
Db 297 RGRGTGRGKPGIN 308

Search completed: April 29, 2003, 09:42:57
Job time : 31 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 09:41:35 ; Search time 15 Seconds
(without alignments)
101.498 Million cell updates/sec

Title: US-09-810-764A-5
Perfect score: 90
Sequence: 1 RGKFKTPGVNIGTIXXV 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	76.7	396	9	US-09-738-626-6937
2	68	75.6	409	10	US-09-845-335-3
3	67	74.4	408	10	US-09-845-335-4
4	66	73.3	394	10	US-09-912-020-247
5	66	73.3	394	10	US-09-845-335-6
6	66	73.3	394	10	US-09-815-242-10362
7	66	73.3	394	10	US-09-815-242-10431
8	66	73.3	394	10	US-09-815-242-11059
9	66	73.3	394	10	US-09-815-242-11059
10	66	73.3	394	10	US-09-815-242-11059
11	62	68.9	22	8	US-08-831-310-9
12	62	68.9	22	10	US-09-488-737-1
13	62	68.9	399	8	US-08-831-310-4
14	62	68.9	399	10	US-09-815-242-11415
15	61	67.8	397	10	US-09-815-242-11995
16	61	67.8	397	10	US-09-815-242-12002
17	57	63.3	398	10	US-09-815-242-13171
18	55	61.1	409	10	US-09-845-335-5
19	53	58.9	394	10	US-09-841-132-491

20	53	58.9	394	10	US-09-841-132-561	Sequence 561, Ap
21	53	58.9	394	12	US-10-007-693-68	Sequence 68, Appl
22	53	58.9	394	12	US-10-007-693-103	Sequence 103, Ap
23	52	57.8	395	10	US-09-815-242-4887	Sequence 4887, Ap
24	52	57.8	395	10	US-09-815-242-10482	Sequence 10482, A
25	48	53.3	394	10	US-09-815-242-5574	Sequence 5574, Ap
26	48	53.3	394	10	US-09-815-242-12442	Sequence 12442, A
27	45	50.0	479	10	US-09-925-300-1515	Sequence 1515, Ap
28	42	46.7	20	9	US-09-736-457-1847	Sequence 1847, Ap
29	42	46.7	20	9	US-09-902-941-1847	Sequence 1847, Ap
30	42	46.7	20	9	US-09-849-626-1847	Sequence 1847, Ap
31	42	46.7	20	9	US-10-017-754-1847	Sequence 1847, Ap
32	42	46.7	200	9	US-09-736-457-789	Sequence 324, Ap
33	42	46.7	200	9	US-09-736-457-789	Sequence 324, Ap
34	42	46.7	200	9	US-09-902-941-324	Sequence 324, Ap
35	42	46.7	200	9	US-09-902-941-789	Sequence 789, Ap
36	42	46.7	200	9	US-09-849-626-324	Sequence 324, Ap
37	42	46.7	200	9	US-09-849-626-789	Sequence 789, Ap
38	42	46.7	200	9	US-10-017-754-324	Sequence 324, Ap
39	42	46.7	200	9	US-10-017-754-789	Sequence 789, Ap
40	42	46.7	201	10	US-09-214-881A-10	Sequence 10, Appl
41	42	46.7	207	9	US-09-736-457-1667	Sequence 1667, Ap
42	42	46.7	207	9	US-09-902-941-1667	Sequence 1667, Ap
43	42	46.7	207	9	US-09-902-941-1913	Sequence 1913, Ap
44	42	46.7	207	9	US-09-849-626-1667	Sequence 1667, Ap
45	42	46.7	207	9	US-09-849-626-1913	Sequence 1913, Ap

ALIGNMENTS

RESULT 1
US-09-738-626-6937
; Sequence 6937, Application US/09738626
; Publication No. US20020197605A1

; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOHI, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6937
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6937

Query Match 76.7%; Score 69; DB 9; Length 396;
Best Local Similarity 73.7%; Pred. No. 0.00054;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGKFKTPGVNIGTIXXV 19

Db 3 KAKFKTKPVNIGTIGHV 21

RESULT 2

US-09-845-335-3

; Sequence 3, Application US/09845335

; Patent No. US20020058266A1

; GENERAL INFORMATION:

; APPLICANT: CLOUGH, BARBARA

; APPLICANT: PREISER, PETER

; APPLICANT: WILSON, ROBERT

; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE

; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS

; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS

; FILE REFERENCE: N68837B GCW PJC DP

; CURRENT APPLICATION NUMBER: US/09/845,335

; CURRENT FILING DATE: 2001-12-26

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/140,466

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-26

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 409

; TYPE: PRT

; ORGANISM: Anacystis nidulans

; US-09-845-335-3

Query Match

Best Local Similarity 75.6%; Score 68; DB 10; Length 409;

Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RGKFKTPKPGVNIQTIXXV 19

| | | | | | | | | | | | | | | |

Db 3 RAKFKTPKPHANIGTIGHV 21

RESULT 3

US-09-845-335-4

; Sequence 4, Application US/09845335

; Patent No. US20020058266A1

; GENERAL INFORMATION:

; APPLICANT: CLOUGH, BARBARA

; APPLICANT: PREISER, PETER

; APPLICANT: WILSON, ROBERT

; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE

; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS

; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS

; FILE REFERENCE: N68837B GCW PJC DP

; CURRENT APPLICATION NUMBER: US/09/845,335

; CURRENT FILING DATE: 2001-12-26

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/140,466

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-26

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 408

; TYPE: PRT

; ORGANISM: Cryptomonas phi

; US-09-845-335-4

Query Match

Best Local Similarity 74.4%; Score 67; DB 10; Length 408;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RGKFKTPKPGVNIQTIXXV 19

| | | | | | | | | | | | | | | |

Db 3 RDKFKSKPHVNIQTIGHV 21

RESULT 4

US-09-912-020-247

; Sequence 247, Application US/09912020

; Patent No. US2002004592A1

; GENERAL INFORMATION:

; APPLICANT: Zyskind, Judith

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION: ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001DVI
; CURRENT APPLICATION NUMBER: US/09/912,020
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/492,709
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/117,405
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 247
; LENGTH: 394
; TYPE: PRT
; ORGANISM: E. Coli
; US-09-912-020-247

Query Match

Best Local Similarity 73.3%; Score 66; DB 10; Length 394;

Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KFERTKPGVNIQTIXXV 19

| | | | | | | | | | | | | | | |

Db 5 KFERTKPHVNIQTIGHV 21

RESULT 5

US-09-845-335-6

; Sequence 6, Application US/09845335

; Patent No. US20020058266A1

; GENERAL INFORMATION:

; APPLICANT: CLOUGH, BARBARA

; APPLICANT: PREISER, PETER

; APPLICANT: WILSON, ROBERT

; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE

; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS

; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS

; FILE REFERENCE: N68837B GCW PJC DP

; CURRENT APPLICATION NUMBER: US/09/845,335

; CURRENT FILING DATE: 2001-12-26

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/140,466

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-26

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 394

; TYPE: PRT

; ORGANISM: Escherichia coli

; US-09-845-335-6

Query Match

Best Local Similarity 73.3%; Score 66; DB 10; Length 394;

Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KFERTKPGVNIQTIXXV 19

| | | | | | | | | | | | | | | |

Db 5 KFERTKPHVNIQTIGHV 21

RESULT 6

US-09-815-242-10362

; Sequence 10362, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10362
LENGTH: 394
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10362

Query Match 73.3%; Score 66; DB 10; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNGTIXXV 19
||||||| ||:|||| |
DB 5 KFERTKPHVNVGTIGHV 21

RESULT 7
US-09-815-242-10431
Sequence 10431, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10362
LENGTH: 394
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10362

NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10431
LENGTH: 394
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10431

Query Match 73.3%; Score 66; DB 10; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNGTIXXV 19
||||||| ||:|||| |
DB 5 KFERTKPHVNVGTIGHV 21

RESULT 8
US-09-815-242-11059
Sequence 11059, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11059
LENGTH: 394
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11059

Query Match 73.3%; Score 66; DB 10; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNGTIXXV 19
||||||| ||:|||| |
DB 5 KFERTKPHVNVGTIGHV 21

RESULT 9
US-09-815-242-11069
Sequence 11069, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIORITY FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11069
LENGTH: 394
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11069

Query Match 73.3%; Score 66; DB 10; Length 394;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIQTIXV 19
||||| |||||
Db 5 KFERTKPHVNVGTIGHV 21

RESULT 10
US-09-815-242-13964
Sequence 13964, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIORITY FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13964
LENGTH: 409
TYPE: PRT
ORGANISM: Salmonella typhi
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(409)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13964

Query Match 73.3%; Score 66; DB 10; Length 409;
Best Local Similarity 76.5%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIQTIXV 19
||||| |||||
Db 20 KFERTKPHVNVGTIGHV 36

RESULT 11
US-08-831-310-9
Sequence 9, Application US/08831310
Patent No. US20020026035A1
GENERAL INFORMATION:
APPLICANT: Kleanthous, Harold et al.
TITLE OF INVENTION: Helicobacter GHPO 1360 and
TITLE OF INVENTION: GHPO 750 Polypeptides and Corresponding Polynucleotides
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/831,310
FILING DATE: 01-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,175
REFERENCE/DOCKET NUMBER: 06132/037001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-08-831-310-9

Query Match 68.9%; Score 62; DB 8; Length 22;
Best Local Similarity 76.5%; Pred. No. 0.00037;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIQTIXV 19
||||| |||||
Db 4 KFERTKPHVNVGTIGHV 20

RESULT 12

US-09-488-737-1
; Sequence 1, Application US/09488737
; Patent No. US20020151462A1
; GENERAL INFORMATION:
; APPLICANT: Lissolo, Ling
; TITLE OF INVENTION: Helicobacter Pylori Membrane Proteins
; FILE REFERENCE: 50019/005002
; CURRENT APPLICATION NUMBER: US/09/488,737
; CURRENT FILING DATE: 2000-01-20
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 08/849,627
; PRIOR FILING DATE: 1996-10-04
; PRIOR APPLICATION NUMBER: FR 95/11,890
; PRIOR FILING DATE: 1995-10-04
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-488-737-1

Query Match 68.9%; Score 62; DB 10; Length 22;
Best Local Similarity 76.5%; Pred. No. 0.00037;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIQTIXXV 19
|| |||| ||||| |
Db 4 KFNRTKPHVNIQTIGHV 20

RESULT 13

US-08-831-310-4
; Sequence 4, Application US/08831310
; Patent No. US20020026035A1
; GENERAL INFORMATION:
; APPLICANT: Kleinhous, Harold et al.
; TITLE OF INVENTION: Helicobacter GHPO 1360 and
; TITLE OF INVENTION: GHPO 750 Polypeptides and Corresponding Polynucleotides
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,310
; FILING DATE: 01-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,175
; REFERENCE/DOCKET NUMBER: 06132/037001
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

; FRAGMENT TYPE: Internal
US-08-831-310-4

Query Match 68.9%; Score 62; DB 8; Length 399;
Best Local Similarity 76.5%; Pred. No. 0.0086;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIQTIXXV 19
|| |||| ||||| |
Db 5 KFNRTKPHVNIQTIGHV 21

RESULT 14

US-09-815-242-11415
; Sequence 11415, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: EUTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11415
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11415

Query Match 68.9%; Score 62; DB 10; Length 399;
Best Local Similarity 76.5%; Pred. No. 0.0086;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 KFERTKPGVNIQTIXXV 19
|| |||| ||||| |
Db 5 KFNRTKPHVNIQTIGHV 21

RESULT 15

US-09-815-242-11995
; Sequence 11995, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

```

; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11995
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11995

```

```

Query Match      67.8%; Score 61; DB 10; Length 397;
Best Local Similarity 70.6%; Pred. No. 0.013;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 3 KEERTKPGVNIQTXXV 19
DQ 5 KPERNKPVNVGTIGHV 21

```

Search completed: April 29, 2003, 09:49:33
Job time : 15 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 05:55:55 ; Search time 5099 Seconds
(without alignments)
9594.394 Million cell updates/sec

Title: US-09-810-764A-6
Perfect score: 1681
Sequence: 1 attcccaataatccccacc.....gttaaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.hig.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1215.4	72.3	138931	2	AP004023	AP004023 Oryza sat
2	1210.6	72.0	2050	8	AF327413	AF327413 Oryza sat
3	1200.6	71.4	1678	8	AF145053	AF145053 Oryza sat
4	790.6	47.0	3079	8	GMNAEFTU	X89058 G.max DNA f
5	787.2	46.8	2313	8	GMTUFA	X66062 G.max tufa
6	774.4	46.0	1584	8	AF234537	AF234537 Pelargon
7	711.8	42.3	1731	8	PSY14561	Y14561 Pisum sativ
8	703.4	41.8	2342	8	TOBTFTU	M94304 Nicotiana t
9	700.2	41.7	2667	8	TOBTUFA2	D11469 Nicotiana s
10	687.2	40.9	2272	8	ATTUFA	X52256 A.thaliana
11	685.6	40.8	1663	8	AF419609	AF419609 Arabidops
12	685.6	40.8	1669	8	AF410329	AF410329 Arabidops
13	685.6	40.8	1688	8	AY074355	AY074355 Arabidops
14	685.6	40.8	109936	8	ATCF9F13	AL080253 Arabidops
15	685.6	40.8	198427	8	ATCHRIV52	AL161552 Arabidops
16	675.2	40.2	1436	8	TOBTUFB	D11376 N. sylvestr
17	675.2	40.2	3675	8	TOBTUFB2	D11470 Nicotiana s
18	658.4	39.2	1390	8	TOBTUFA	D11375 N. sylvestr
19	630.2	37.5	298750	1	AP005375	AP005375 Thermosyn
20	629.8	37.5	1337	8	GMTUFB2	Y15108 Glycine max
21	622.2	37.0	5139	1	ANORF150	X17442 A.nidulans
22	604.4	36.0	10492	1	AE001892	AE001892 Deinococc
23	604.2	35.9	9997	1	AE002041	AE002041 Deinococc
24	590.4	35.1	348550	1	AP003596	AP003596 Nostoc sp
25	565	33.6	143308	1	D90913	D90913 Synecocyst
26	557.6	33.2	1741	1	AF007125	AF007125 Streptomy
27	554.4	33.0	5273	1	SPSTR	X15646 Spirulina p
28	553.4	32.9	200799	8	AF137379	AF137379 Nephrosel
29	548.4	32.6	1358	1	S79408	S79408 tufl-elonga
30	548.4	32.6	23087	1	SCD40A	AL161691 Streptomy
31	547	32.5	1900	1	SCTUFLFUS	X77039 S.coelicolo
32	539.8	32.1	2012	1	SATUFBRNA	X82820 S.aurantia
33	538	32.0	1191	1	AB073986	AB073986 Myxococcu
34	538	32.0	1826	1	PRU67308	U67308 Planobispor
35	536.4	31.9	1191	1	TCHOESTUF	X76871 T.cuprinus
36	536.4	31.9	2742	1	PRFUSTUF	X98830 P.rosea fus
37	535.8	31.9	2250	1	AF153618	AF153618 Streptomy
38	533.6	31.7	1374	1	AF368284	AF368284 Streptomy
39	532.6	31.7	9150	1	AE011727	AE011727 Xanthomon
40	531.8	31.6	5203	1	TCU78300	U78300 Thiobacillu
41	531.8	31.6	13867	1	AE011726	AE011726 Xanthomon
42	529.8	31.5	13486	1	AE004842	AE004842 Pseudomon
43	529.4	31.5	2836	1	SRTUFL	X57057 S.ramoclassi
44	526.4	31.3	11171	1	AE004843	AE004843 Pseudomon
45	524.8	31.2	7934	1	AE005800	AE005800 Caulobact

ALIGNMENTS

RESULT 1	AP004023	138931 bp	DNA	linear	HTG 21-MAR-2002
LOCUS	Oryza sativa (japonica cultivar-group)	chromosome 2 clone			
DEFINITION	OJ1126_D09, *** SEQUENCING IN PROGRESS ***				
ACCESSION	AP004023.1	GI:15130685			
VERSION	HTG: HTGS_PHASE2				
KEYWORDS	Oryza sativa (japonica cultivar-group)	(cultivar:Nipponbare) DNA,			
SOURCE	Clone:OJ1126_D09.				
ORGANISM	Oryza sativa (japonica cultivar-group)				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
	Ehrhartoideae; Oryzeae; Oryza.				
REFERENCE	1				

AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC clone: OJ1126.D09
 JOURNAL Published Only in Database (2001)
 REFERENCE 2 (bases 1 to 138931)
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Direct Submission
 JOURNAL Submitted (08-AUG-2001) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@kenias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 COMMENT The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 source Location/Qualifiers
 1. 138931
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OJ1126.D09"
 BASE COUNT 39463 a 30336 c 30692 g 37874 t 566 others
 ORIGIN
 Query Match 72.3%; Score 1215.4; DB 2; Length 138931;
 Best Local Similarity 84.4%; Pred. No. 3e-154;
 Matches 1405; Conservative 0; Mismatches 251; Indels 9; Gaps 3;
 QY 4 CCAAAATATCCACCTCCGCTGCTGCTCCGCGCCGCCATGCGCTCCCTCACTCG 63
 DB 98236 CCCCCCACCCTCTCTGCGGCTCCCTCTCCGCGGCTCCGCCATGCGCTCCCTCG 98295
 QY 64 GCGTCCACTT---CACTTCTTCCCGCAGGCTCTCTCACTCAGGAGCGGCATCCCTCTC 120
 DB 98296 GCCTCCGATACACCTGCTGCTGCTTCTCCACCTCTCTTCCAGCGCGCTCGCTCC 98355
 QY 121 TCCACCCCTGGCTTCTCCGCGACGCTTGCCTGCTGCGGCTCGGAGCCAGGCGCGCAGTG 180
 DB 98356 TCCGTCGGATTCTCTCGCGCGCGGCTTCCGCGCAGCGCGCGCGGCTCCAGG 98415
 QY 181 GCGCGCGCGCGCGCGCGCTGCTGCTGCTGCGCGCGCGCGAGGCGCAAGTTCGAGCGC 240
 DB 98416 GGCACGGCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 98475
 QY 241 ACCAAACACACGCTCAACATAGGACCATCGGCTATGTCAGCACACCAAGAACCACTCTC 300
 DB 98476 ACCAAGCGCACGTCACATCGGACCATCGGCTATGTCAGCACACCAAGAACCACTCTC 98535
 QY 301 ACCGCGCGCTACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 DB 98536 ACCGCGCGCTACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 98595
 QY 361 GAGATCACCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 DB 98596 GAGATCACCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 98655
 QY 421 GAGTACGAGACCGGACCGGCTACGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 98656 GAGTACGAGACCGGACCGGCTACGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 98715
 QY 481 GTCAAGNATATGATACCGCGCTGCGAGATGGAGCGGTGCCATTCCTGCTGATTCGGGT 540
 DB 98716 GTCAAGNATATGATACCGCGCTGCGAGATGGAGCGGTGCCATTCCTGCTGATTCGGGT 98775

QY 541 GCCGAGCGGCCCATCGCGCAGACCAAAAGAGACACATCTCTCTGCCAAGCAAGTCGGTGT 600
 DB 98776 GCCGAGCGGCCCATCGCGCAGACCAAAAGAGACACATCTCTCTGCCAAGCAAGTCGGTGT 98835
 QY 601 CCCAAGATCGTGTCTTCTCTCAACAAGAGGACATGCTGCGAGCAGGAGGAGTCTCGAG 660
 DB 98836 CCCAAGATGTTGTCTTCTCTCAACAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 98895
 QY 661 CTCGTCGAGCTCGAGTCCGCGAGCTGCTACGAACTACGAGTACGAGCGGACGACGACGTA 720
 DB 98895 CTCGTCGAGCTCGAGTCCGCGAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 98955
 QY 721 CCAATCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 DB 98956 CCAATCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 99015
 QY 781 AAGCGGGGCGAGATGAGTGGTGCATACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
 DB 99016 AAGCGGGGCGAGATGAGTGGTGCATACATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 99075
 QY 841 TATATTCAGTCCCGCAGAGGCGAGCTGACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 99076 TATATTCAGTCCCGCAGAGGCGAGCTGACCTCCGCTGCTGCTGCTGCTGCTGCTGCT 99135
 QY 901 TTCTCCATCACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 DB 99136 TTCTCCATCACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 99195
 QY 961 ATTGGTACACAGTTCGATATCGTCGGAATCGGGAATCGGGAATCGGGAATCGGGAATCG 1020
 DB 99196 GTTGGGAGACGCTGATATCGTCGCTATCGGGAATCGGGAATCGGGAATCGGGAATCG 99255
 QY 1021 GTTGAGATGTTCCAGAAGACCATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 DB 99256 GTTGAGATGTTCCAGAAGACCATGATGATGATGATGATGATGATGATGATGATGATG 99315
 QY 1081 CTCCTGCTATGACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 DB 99316 CTCCTGCTATGACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 99375
 QY 1141 ATCACACGCGACACCAAGTTCGAGCTGCTGATGATGATGATGATGATGATGATGATGATG 1200
 DB 99376 ATCACACGCGACACCAAGTTCGAGCTGCTGATGATGATGATGATGATGATGATGATG 99435
 QY 1201 CGCACTCACCTTCTTCCCTGCTTACCGCCACAGATTCATGCTGCGGACCACTGATGTG 1260
 DB 99436 CGCACTCACCTTCTTCCCTGCTTACCGCCCTCAGTTCATGCTGCGGACTACCGATGTG 99495
 QY 1261 ACAGGAGTGTGACTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 DB 99496 ACAGGAGTGTGACTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 99555
 QY 1321 GGTGACCTGATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
 DB 99556 GGTGACCTGATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 99615
 QY 1381 AGTTTGTATCCGCTGAGGCTGTTAAGACCTGTTGTCGCGGTGTCATCAACAAATCAT 1440
 DB 99616 AGTTTGTATCCGCTGAGGCTGTTAAGACCTGTTGTCGCGGTGTCATCAACAAATCAT 99675
 QY 1441 GAGTAACTGGATATACATATCCACCATGAGATTTTCTCTGTTTCTACTCAACGGGACAT 1500
 DB 99676 GAGTAACTGGATATACATATCCACCATGAGATTTTCTCTGTTTCTACTCAACGGGACAT 99735
 QY 1501 GCTCGTGTGTTTATGTTGCTGAGTTTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 DB 99736 GCTCGTGTGTTTATGTTGCTGAGTTTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 99795
 QY 1561 ACATTTTTTTTTGTCAAGTGAATTTGATGAATTTATGACATTCACGACAAAGATTCACA 1620
 DB 99796 GCATTTTTTCTCT--TCAAGTGAATTTGATGAATTTATGACATTCACGACAAAGAT----A 99849

QY 1321 GGTGACCGTATCAAAATGATTTGCTAGCTCATCCAGCCCTGTTGCTGTGTGAGCAGGGTATG 1380
 Db 1387 GGTGACCGTGTCAAGATGTTGTGGAGCTCATCCAGCCCGTCTGTTGTGAGCAGGAATG 1446
 QY 1381 AGTTTCTATCCGTGAGGGTGTAAAGACCGTGTGGTGGCGGTGTCTATCAACAAATCAT 1440
 Db 1447 AGTTTCTATCCGTGAGGGTGTAAAGACCGTGTGGTGGCGGTGTCTATCAACAAATCAT 1506
 QY 1441 GAGTAAATCTGATATACATATACCATCATGAGAAATTTTCTGTTTACTCAAGAGCAT 1500
 Db 1507 AAGTAAATCTGATGAGGATATACCATCATGAGAAATTTTCTGTTTACTCAAGAGCAT 1566
 QY 1501 GCTCCGTAGTGTATATGTTGTTAGGTTTATAGGGTGTCTCATGTCGAATTTGATATG 1560
 Db 1567 GCTCTGTAGTGTATATGTTGTTAGGTTTATAGGGTGTCTCATGTCGAATTTGATATG 1626
 QY 1561 ACATTTTCTGTTGTTCAAGTAAATTTGTCATATTTATGATATTCAGCAAAAGATTCACA 1620
 Db 1627 GCATTTTCTCT--TCAAGTAAATTTGTCATATTTGTTGATATTCAGCAAAAGAT--A 1680
 QY 1621 TATCTGTTGCAACTCATTTGGCTAAGAGGTGCATCTACTGTGA 1665
 Db 1681 CATCTGTTGCAACTATTTGGCTATGCTGTGCATCTACTGTCA 1725

AF145053 1678 bp mRNA linear PLN 06-DEC-1999
 Oryza sativa chloroplast translational elongation factor Tu (tufa)
 mRNA, complete cds; nuclear gene for chloroplast product.

AF145053
 AF145053.1 GI:6525064

Oryza sativa (japonica cultivar-group).
 Oryza sativa (japonica cultivar-group).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 Lee, J.H., Lee, J.W., Chung, Y.Y., Paek, K.H., Shin, J.S., Yun, C.H. and Kim, J.K.
 Cloning and characterization of the chloroplast elongation factor EF-Tu cDNA of *Oryza sativa* L.
 Mol. Cells 9 (5), 484-490 (1999)
 20062172
 10597036

REFERENCE 2 (bases 1 to 1678)
 Lee, J.H., Lee, J.W., Chung, Y.Y., Paek, K.H., Shin, J.S., Yun, C.H. and Kim, J.K.
 Direct Submission
 Submitted (22-APR-1999) Biology, Korea University, Seong Buk Gu, Anam-Dong, 5-1, Seoul, South Korea

FEATURES
 Location/Qualifiers
 1..1678
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="japonica Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="1"
 1..1678
 /gene="tufa"
 62..1465
 /gene="tufa"
 /codon_start=1
 /product="chloroplast translational elongation factor Tu"
 /protein_id="AA15312.1"
 /db_xref="GI:6525065"
 /translation="MASLASASTLSVSTSSKPRLGSSVGFSSPARPRTAAAA
 SKGTGRRAGLLVMAARKFKERTPHVNIQTHIGHVDHGKTLTLAALFWLASVGGSP
 KKYDEIDAPERARGITINTVEYEFETRIYAHVDCPHADIVKMMITGAAQMDGA
 ILVSGADPMPPQKEHILLAKQVVPKIVVFLNKKQDVDEELQLLEVEVRELLSS
 YVGDGEPIVAGSALKALEMANPAIKRGDWDVDFISLIDVDNYIPVPOQRTD
 LPFLIAEDVFSITGRGTATGRIERTGVKVDVIVIGRETRNCTVTGVEMFOKTM
 DDAMAGDNVLLRGMOKDDIERGMVLAKPASITPHTKFDVVVVLKDKSGRHSPPF
 PCYRPOFYRTTDTVGNVYPKIMNDKDEAKMCMFGDRVKVVELLIQPVACEQGMRAI

BASE COUNT 326 a 512 c 485 g 355 t
 ORIGIN
 Query Match 71.4%; Score 1200.6; DB 8; Length 1678;
 Best Local Similarity 83.9%; Pred. No. 5.2e-152;
 Matches 1395; Conservative 0; Mismatches 259; Indels 9; Gaps 3;

QY 4 CCCAAATAATCCCAACCTCCCGTCTGCTCCGCGCGCCGCGCATGGCTCCCTCAACCTCG 63
 Db 20 CCCCCCACTCTCTGCGGCTCCCTCTCCGCGGCTCCGCGCATGGCTCCCTCCGCTCC 79
 QY 64 GCGTCCACTT---CACTCTCTTCCCGCAGCCTCTCTATCCAGGAGCGCATCCGCTCTC 120
 Db 80 GCCTCCGATCATCACTCCCTGCTCTCTCCACCTCTCTCTCAAGCGCGCTCGGCTCC 139
 QY 121 TCCACCCCTCTGGGTTCTCCGCGAGCTCGCGGCTGCGGAGCCAGGGGGGAGTGC 180
 Db 140 TCCGTCCGATTTCTCTCGCGCGCGGTTCCGGCGCACGGCGCGGCGGCTCCAAAG 199
 QY 181 GCGCGCGCGCGCGCGGCTCTGCTGCTGCTGCGCGCGCGGAGGGGCAAGTTTCAGCGC 240
 Db 200 GCGAGGGCGCGCGCGGCTCTGCTGTAATGCGCGGCGGAGGGGAGTTTCGACGG 259
 QY 241 ACCAAACACACGTCAACATAGGCACCATCGGCCATGTGACACACGAGAAAGACACTCTC 300
 Db 260 ACCAAGCGCACGTCAACATCGGCACCATCGGCCACGTGACACACGAGAAAGTACGCTG 319
 QY 301 ACCGCGCGCTCACCATGGTGTCTCGCTCGTTCGTTGCGAGCGGCTTAAGAAGTAGCAC 360
 Db 320 ACGCGCGCTCACCATGGTGTCTCGCTCGTTCGCGGCGGAGCGGCCCAAGAAAGTAGCAC 379
 QY 361 GAGATCGACGCGCGCGCGGAGCGCGGCTATCACCATCAACACCGCACCGCTC 420
 Db 380 GAGATCGACGCGCGCGGAGCGCGGCTATCACCATCAACACCGCACCGCTC 439
 QY 421 GAGTACGAGACCGAGACCGCCACTACGACACGTGAGTCCGCCGCCGCGGAGTAT 480
 Db 440 GAGTACGAGACCGAGACCGCCACTACGACACGTGAGTCCGCCGCCGCGGAGTAT 499
 QY 481 GTCAAGATATGATCACGCGCTGCGCATGAGCGGTCATCTCTGTCGTATCCGGT 540
 Db 500 GTCAAGATATGATCACGCGCTGCGCATGAGCGGTCATCTCTGTCGTATCCGGT 559
 QY 541 GCGGACGCGCCATCGCGCAGACCAAGACATCTCTCGCCCAAGCAAGTGGTGT 600
 Db 560 GCGGACGCGCCATCGCGCAGACCAAGACATCTCTCGCCCAAGCAAGTGGTGT 619
 QY 601 CCAAGATCGTGTCTCTCAACAAGACATGTGTCAGCAGAGGAGTGTCTCGAG 660
 Db 620 CCAAGATCGTGTCTCTCAACAAGACATGTGTCAGCAGAGGAGTGTCTCGAG 679
 QY 661 CTCGTGAGCTCGAGTCCGCGAGCTCTCAGCAACTACGAGTACGACGCGCAGCGTA 720
 Db 680 CTCGTGAGCTCGAGTCCGCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 739
 QY 721 CCAATCGTGTGCTCCGCCCTCAAGCGCTCGAGGCTCTCATGTGTCAACCTCGCTTG 780
 Db 740 CCAATCGTGTGCTCCGCCCTCAAGCGCTCGAGGCTCTCTCTCTCTCTCTCTCTCT 799
 QY 781 AAGCGCGCGCAGTATGAGTGGTCTGACTATCTTCTCTCTCTCTCTCTCTCTCTCT 840
 Db 800 AAGCGCGCGCAGTATGAGTGGTGGAGGATCTTCTCTCTCTCTCTCTCTCTCTCT 859
 QY 841 TATATTCCAGTCCCGCAGAGCAGCTGACTCTCCCGTCTTCTCTCTCTCTCTCTCT 900
 Db 860 TATATTCCCTGTCCACAGCGCAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 919
 QY 901 TTCTCCATCACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 Db 920 TTCTCCATCACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 979
 QY 961 ATTGGTACACAGTGCATATCTCTCGGAAATCGGGGACACCCGGAACTCGACGGTCACTGGT 1020

QY 673 GAGTCCGCGAGCTGCTCAGCACTACGAGTACGACGGCAGCAGGTACCAATCGTCGCT 732
 Db 1885 GAGTCCGCGAGCTCTTCCAGTACGAGTCCCGGAGCAGCATGTCGGATATCTCC 1944
 QY 733 GGCTCCGCGCTCAAGCGCTCGAGGCTCTCATGCTCAACCCCTGCCCTTGAAGCGCGCGAC 792
 Db 1945 GGTTCGCGCTTGTCTTGTGAAGCTTTGATGCCAACCGCTCCATCAAGCGCGGGAA 2004
 QY 793 GATCAGTGGGTGCGACTACATCTCTCGTGGTGGTGAAGAGTGAATCCCTATATCCAGTC 852
 Db 2005 AACCAATGGGTGATAAAATTTACGAGCTCATGGAAGCTGTGATGACTACATCCCATC 2064
 QY 853 CCGCAGGAGCAGTACCTCCCGTCTTCTGCTGGTGAAGAGTGTCTCTCCATCAC 912
 Db 2065 CCTCAGCGCCAACTGAATCCCTTTCTGCTGCCATGAGAGTGTTCACCATCAC 2124
 QY 913 GGTGCTGGTACAGTGGCCACTGCGCTATAGAGCTGGCAGCTCAAGATTTGGTGACACA 972
 Db 2125 GGTGCTGGAACCGTCGCCACCGGAAGGTTGAGAGAGTACTATTAGGTTGGAGAGCT 2184
 QY 973 GTCGATATCTCGGAATCCCGGACACCGGAAGTGCAGGCTACTGGTGTGAGATGTC 1032
 Db 2185 GTTGATATGTTGGTGTGAAGGACACTAGGAACACCACTGTGACTGGGTGGAAATGTTT 2244
 QY 1033 CAGAAGACCATGGATGATGATGCGCGGAGACAATGTTGGGCTGCTCGTGGTATG 1092
 Db 2245 CAGAGATTTCTGATGAGGCTTTGGCTGGGACAATGTTGGCTGTAGGGGTATT 2304
 QY 1093 CAGAAGATGACATTAAGAGAGCATGCTGCTGGCAAGCGTGGCTCTATCACACCGCAC 1152
 Db 2305 CAGAAAATGATATTCAGAGAGGATGTTTGGCTAAGCTTGAACCATTTACTCCTCAC 2364
 QY 1153 ACCAAGTTTTCAGGCTGTTGTATGCTTTAAGAAGGAAGAGGTTGGCCGACACTCACT 1212
 Db 2365 ACCAAGTTTTCGCAATGTTGTATGCTTTGAAGAAGGAAGAGGTTGGGAGGCACCTCC 2424
 QY 1213 TTCTTCCCTGTTACCGCCACAGTCTACATGCGGACACACTGATGACAGGAGTGTG 1272
 Db 2425 TTCTTTCCGCTACCGGCTCAGTCTACATGAGGACACTACCATGTTTACTGGCAAGTA 2484
 QY 1273 ACTGATATTAATGACAAAGATGAGGAGGAGATGTCATGTCGCTGGTGACCGTATC 1332
 Db 2485 ACTGAAATTAATGACAAAGATGAGAGTCCAAGATGTTATGCCGGCGACCGTGT 2544
 QY 1333 AAAATGATTTGAGCTATCCAGCTGTTGCTGTGACAGGATGATGAGTTTGTCTATC 1392
 Db 2545 AAGCTGGTGTGAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 2604
 QY 1393 CGTGGGGTGAAGACCGTTGTCGCGGTGTCATCAACAAATCATTTAGTA 1445
 Db 2605 AGAGAAGGAGGCAAACTGTGGAGCTGGTGTATCCAAATCCATCATTTAGTA 2657

RESULT 5
 GMTUFA
 LOCUS
 DEFINITION G. max tufa gene for chloroplast translation elongation factor
 EF-Tu.
 X66062
 VERSION X66062.1 GI:18775
 KEYWORDS translation elongation factor; tufa gene.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 2313)
 Stutz.E.
 Direct Submission
 Submitted (17-MAY-1992) E. Stutz, Lab. de Biochimie Vegetale,
 Universite de Neuchatel, Chantemerle 18, CH-2000 Neuchatel,
 SWITZERLAND

REFERENCE 2 (bases 1 to 2313)
 AUTHORS Bonny,C. and Stutz,E.
 TITLE Soybean (glycine max l.) nuclear DNA contains four tuf genes
 coding for the chloroplast specific translation elongation factor
 EF-Tu
 JOURNAL Chimia 47, 247-249 (1993)
 FEATURES Location/Qualifiers
 source
 1..2313
 /organism="Glycine max"
 /cultivar="Maple Arrow"
 /db_xref="taxon:3847"
 /tissue_type="leaves"
 /clone_lib="genomic dna"
 220..253
 /note="light responsive element"
 misc_feature
 611..617
 TATA_signal
 gene 710..2149
 /gene="tufa"
 CDS 710..2149
 /gene="tufa"
 /codon_start=1
 /product="EF-Tu"
 /protein_id="CAA46864.1"
 /db_xref="GI:18775"
 /translation="MAVSSATASSKLILLPHASSSSLSNSTPFRSTTHKLTPLSS
 SFHPTVLRTPSTSTTPRTTVAARGKFKPKPHVNIPTIGVHDKGKTLTAAL
 TMAALGNSAPKKYDEIDAAPEERARGITINTATVEYETENRHYAHVDCPGHADYVK
 NMTVGAQMGGALLVSGADGMPQPKHEIILAKOVGPNVYVFLNKQDQVDBELLQ
 LVEIEVDLLSSFEFGDDPTVSGSALLALEMANPAIKRDNEDWVKIFOLMDEV
 DNYIPIRQTDLPFLAVEDVFSITGRGTATGRVETIKYGETVDLVGLRETRNT
 TTVGVEMFQKIDLEALAGDNVLLRQKTDIQRGMVLAKPGTITPHFKFSAIVTVL
 KKEGGHSPFFAGYRPFYMRMTDVTGKVTISMDKDEESTVLPQDRVKMVELIV
 PVACEQGMFAIREGGKTVGAGVIQSIIE"
 transit_peptide 710..922
 /gene="tufa"
 mat_peptide 923..2146
 /gene="tufa"
 polyA_site 2312
 /product="EF-Tu"
 BASE COUNT 675 a 609 c 439 g 590 t
 ORIGIN
 Query Match 46.8%; Score 787.2; DB 8; Length 2313;
 Best Local Similarity 72.4%; Pred. No. 1.5e-96;
 Matches 1020; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
 QY 53 CCCTCAGCTCGGCTCCACTTCCTTCCCGAGGCTCTCTATCCAGAGCGCGCA 112
 Db 756 CCCATGATCATCATCTCTTCCCTTAATCCACACCTTCCGCTCTCCACCAACA 815
 QY 113 TCCGCTCTCCACCCCGCTTCTCCGCGCAGCCTGCGCGGCTCGGAGCCAGGGG 172
 Db 816 CCCACAAATACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 875
 QY 173 CGGAGTGGGCGCGGCGGCGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 232
 Db 876 CCCCTTCTCCACCAACCCCGCTGCGACCTTACCGTCCGCGCGCGCGCGCAAT 935
 QY 233 TCGAGCGCACCAACACACGTCACATAGCACCATCGCCCATGTCGACACCGGAAGA 292
 Db 936 TCGAGCGCAAGACCCCGCTGAACATCGGCACCATCGGCCACGTGGACCGGCAAGA 995
 QY 293 CCAGCTTCACGGCGCGCTCACCATGCTGCTCGCTCCGCTGGGCGAGCGCGCTAAGA 352
 Db 996 CCACCTTACCGCTGCGCTGACCATGCTCTTCCGCGCTAGGCAACAGCGCCCCAAA 1055
 QY 353 AGTACGAGAGATCGAGCGCGCGCGCGAGAGCGCGCGCGGTATCACATCAACACCG 412
 Db 1056 AATACGAGAGATCGAGCGCGCGCGCGAGAGCGCGCGCGCATCACCATCAACACCG 1115
 QY 413 CCACCGTCGAGTACGAGACCGGACCGCGCGCTACGACACGTCGACTGCCCGCGCACG 472

[illegible]


```
Db 827 ATTACCGGTGCTGCCAAATGAGCGGGCTATTCTTCTGTTTTCAGGTGCTGATGGTCC 886
Qy 553 ATGCGCAGACAAAGAGCAGCATCTCTCGCCAAAGCAAGTCGGTGTCCCAAGATCGTT 612
Db 887 ATGCCACAGACTAGGAACACATTTTCTGCTAAGCAAGTAGTGTCGCCACACATGGTT 946
Qy 613 GTCTTCTCTCAACAAGAGACATCTCGAGCAGCAGGAGCTGCTCGAGCTCGTCGAGCTC 672
Db 947 GTGTTCTTGAACAACAAGATCAGTTGATGATGAGGAGCTGCTCAGCTTGTGAGTTG 1006
Qy 673 GAGTCCCGAGCTGCTCAGCAACTAGAGTACAGCGGAGCAGCTACCAATCGTCGCT 732
Db 1007 GAGTTAGAGAGTATTATGTCAGATTTAGAGTTCTCGTGTATGATATTCCTATTATTTCT 1066
Qy 733 GGCTCCGCTCAAGGCGCTCGAGGCTCTCATGCTCAACCTGCTGCTGAAGCGCGGAC 792
Db 1067 GGCTCTGCTCTTTGGCTTTAGAGGCTTTGATGGCTAATCCTAGTATTAAGAGAGTGAA 1126
Qy 793 GATGAGTGGTGCAGTACATCTTCTGCTGTTGATGAAGTGGATTCCTATATTTCCAGTC 852
Db 1127 AATCAATGGGTGATAAGATATATGATGATGAGGAGCTGTTGATAGTTATATTTCTATT 1186
Qy 853 CCGCAGAGGAGAGTACCTCCCTCTTCTGCTGCTGTTGAAGATGCTTCTCCATCACC 912
Db 1187 CCAGTTAGGCAAACTGAATTCCTCTTCTGATGGCTATTAAGATGTTGTTTCCATTACT 1246
Qy 913 GGTGCTGTACAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
Db 1247 GGTAGAGTACTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1306
Qy 973 GTCGATATCTGCGAATCCGGGACACCCGCAACTGCAAGTCACTGCTGTTGAGATGTTTC 1032
Db 1307 GTTGATATTTGAGTTTGAAGGAGACTAGGAGTACTACCGTGAAGGCTGTTGAGATGTTT 1366
Qy 1033 CAGAAGACCATGATGATGCTGCGGAGCAATGCTGCGTCTGCTGCTGCTGCTGCTGCTGCTG 1092
Db 1367 CAAGAAGATTTGATGAAGCAATGCTGCGGAGCAATGCTGCGTCTGCTGCTGCTGCTGCTG 1426
Qy 1093 CAGAAGGATGATGATGAAGAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1152
Db 1427 CAAAAGATTTGATATTCAGAGAGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1486
Qy 1153 ACCAAGTTGAGGCTGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1212
Db 1487 ACCAAGTTGAGGCTGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1546
Qy 1213 TTCTTCTCTGCTTACCGCCACAGTCTTACATGCGGCAACTGATGATGACAGGAGTGTG 1272
Db 1547 TTCCTTTCAGGCTACAGGCTCAGTTTACATGAGGACTACTGATGACCGGAAGGTT 1606
Qy 1273 ACTAGGATTAATGACAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1332
Db 1607 ACTTCCATTAGGCTGATGAAGGAGAGGAATTAAGATGCTGCTGCTGCTGCTGCTGCTGCTG 1666
Qy 1333 AAAATGATGTTGACGCTCATCCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
Db 1667 AACTTGGTGTGAGCTCATTTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1726
Qy 1393 CGTGAGGTTGTAAGACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1452
Db 1727 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1786
Qy 1453 TATAATATATCA 1465
Db 1787 AATAGCGGAGCAA 1799

RESULT 9
TOBTUFA2
LOCUS
DEFINITION Nicotiana sylvestris tufa gene for chloroplast elongation factor
Tua, complete cds.
ACCESSION D11469
```

D11469.1 GI:459238
g-binding protein; elongation factor; nuclear-encoded chloroplast
elongation factor TuA(EF-TuA).
Nicotiana sylvestris DNA, clone_lib:lambda Dash.
Nicotiana sylvestris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 2667)
Sugita,M., Murayama,Y. and Sugiyama,M.
Structure and differential expression of two distinct genes
encoding the chloroplast elongation factor Tu in tobacco
Curr. Genet. 25 (2), 164-168 (1994)
94373864
2 (bases 1 to 2667)
Sugiyama,M.
Direct Submission
Submitted (19-JUN-1992) Masahiro Sugiyama, Nagoya University, Center
for Gene Research (Tel:052-789-3081, Fax:052-789-3081)
Location/Qualifiers
1. .2667
/organism="Nicotiana sylvestris"
/db_xref="taxon:4096"
/clone_lib="lambda Dash"
257..265
CAAT_signal
288..294
TATA_signal
331..337
TATA_signal
352..357
misc_feature
388
/note="transcription start site"
551
/note="transcription start site"
668..2104
/gene="tufa"
668..2104
/gene="tufa"
/codon_start=1
/product="chloroplast elongation factor TuA(EF-TuA)"
/protein_id="BAA0207.1"
/db_xref="GI:459239"
/translation="MASTSAATSTTKLVSSNSTNPLPSSTKPSKLILSSSTPNC
STLHSPATPSTATHRRRTVRAARGKPKPHVNIQTIGVPHGKTLTAALT
MALASMGSPAKYDEIDDAPEERARGITINIAIVEYETENRHHVHDCPGHADYVN
MITGAAMDGAIVCSGADPMPQTKHILLAKOVGVNMFVFNKQDQVDEELLQ
VELEVRELLSYEPFGDDIPTISGALLAALMANPSIKRGENQWDKIVELMDVD
SYIPIPVROTLPFLMAIEDVTSITGRTVATGRTVGTGRTVGTGRTVGTGRTVGT
VTGEMFOKILDEAMAGNVGLLRGIQKIDIORGMVLAKPGTITPTKFAIVVILK
KEGGRHSPPESGYRPOFYMTDVTGKVTISITTDKGEESKMWMPGDRVNLVVELIMP
VACGGHRFAIRBGGKTVGAGVIQKIE"
transit_peptide 668..877
/gene="tufa"
mat_peptide 878..2101
/gene="tufa"
/product="chloroplast elongation factor TuA(EF-TuA)"
BASE COUNT 758 a 528 c 533 g 848 t
ORIGIN
Query Match 41.7%; Score 700.2; DB 8; Length 2667;
Best Local Similarity 71.9%; Pred. No. 7e-85;
Matches 915; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
Qy 193 CGCGGGCGCTGCTGGTGGTGGCGGGGAGGGGCAAGTTGCGAGCGCAACACAC 252
Db 851 CGCCACCGCGCTTTTACTGCGCGCTGCGCGTGGCAAATTCGAGCGGAAAAACCCCAT 910
Qy 253 GTCAATAGACACCATCGCCATGTCGACCAACGAAAGACCACTCTCACCGCCGCTC 312
Db 911 GTCAATAGACACCATGTCGCGCTGTCGCGTGGCAAGAACTACTCTCACCGCTCTTTA 970
Qy 313 ACCATGGTCTGCTCGCTCGCTGCGCGCGCGCTTAAGAGTACGAGAGATCGACGCC 372
Db 971 ACCATGGTCTGCTCTATGCGAAATCCCGCCCTTAAGAAATAGACGAAATTCATGCC 1030

827 ATTACCGGTGCTGCCAAATGAGCGGGCTATTCTTCTGTTTTCAGGTGCTGATGGTCC 886
553 ATGCGCAGACAAAGAGCAGCATCTCTCGCCAAAGCAAGTCGGTGTCCCAAGATCGTT 612
887 ATGCCACAGACTAGGAACACATTTTCTGCTAAGCAAGTAGTGTCGCCACACATGGTT 946
613 GTCTTCTCTCAACAAGAGACATCTCGAGCAGCAGGAGCTGCTCGAGCTCGTCGAGCTC 672
947 GTGTTCTTGAACAACAAGATCAGTTGATGATGAGGAGCTGCTCAGCTTGTGAGTTG 1006
673 GAGTCCCGAGCTGCTCAGCAACTAGAGTACAGCGGAGCAGCTACCAATCGTCGCT 732
1007 GAGTTAGAGAGTATTATGTCAGATTTAGAGTTCTCGTGTATGATATTCCTATTATTTCT 1066
733 GGCTCCGCTCAAGGCGCTCGAGGCTCTCATGCTCAACCTGCTGCTGAAGCGCGGAC 792
1067 GGCTCTGCTCTTTGGCTTTAGAGGCTTTGATGGCTAATCCTAGTATTAAGAGAGTGAA 1126
793 GATGAGTGGTGCAGTACATCTTCTGCTGTTGATGAAGTGGATTCCTATATTTCCAGTC 852
1127 AATCAATGGGTGATAAGATATATGATGATGAGGAGCTGTTGATAGTTATATTTCTATT 1186
853 CCGCAGAGGAGAGTACCTCCCTCTTCTGCTGCTGTTGAAGATGCTTCTCCATCACC 912
1187 CCAGTTAGGCAAACTGAATTCCTCTTCTGATGGCTATTAAGATGTTGTTTCCATTACT 1246
913 GGTGCTGTACAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
1247 GGTAGAGTACTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1306
973 GTCGATATCTGCGAATCCGGGACACCCGCAACTGCAAGTCACTGCTGTTGAGATGTTTC 1032
1307 GTTGATATTTGAGTTTGAAGGAGACTAGGAGTACTACCGTGAAGGCTGTTGAGATGTTT 1366
1033 CAGAAGACCATGATGATGCTGCGGAGCAATGCTGCGTCTGCTGCTGCTGCTGCTGCTGCTG 1092
1367 CAAGAAGATTTGATGAAGCAATGCTGCGGAGCAATGCTGCGTCTGCTGCTGCTGCTGCTG 1426
1093 CAGAAGGATGATGATGAAGAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1152
1427 CAAAAGATTTGATATTCAGAGAGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1486
1153 ACCAAGTTGAGGCTGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1212
1487 ACCAAGTTGAGGCTGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1546
1213 TTCTTCTCTGCTTACCGCCACAGTCTTACATGCGGCAACTGATGATGACAGGAGTGTG 1272
1547 TTCCTTTCAGGCTACAGGCTCAGTTTACATGAGGACTACTGATGACCGGAAGGTT 1606
1273 ACTAGGATTAATGACAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1332
1607 ACTTCCATTAGGCTGATGAAGGAGAGGAATTAAGATGCTGCTGCTGCTGCTGCTGCTGCTG 1666
1333 AAAATGATGTTGACGCTCATCCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
1667 AACTTGGTGTGAGCTCATTTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1726
1393 CGTGAGGTTGTAAGACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1452
1727 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1786
1453 TATAATATATCA 1465
1787 AATAGCGGAGCAA 1799

RESULT 9
TOBTUFA2
LOCUS
DEFINITION Nicotiana sylvestris tufa gene for chloroplast elongation factor
Tua, complete cds.
ACCESSION D11469


```
QY 373 GCCCGGAGGCGCGCGCGGTATCACCATCAACACCGCACCGTGGAGTACGAGACC 432
|||||
Db 1031 GCCCAGAGAGCGTCTCGTGTATTAATAACACTGCTACTGTTGAGTACGAGACT 1090
|||||

QY 433 GAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 492
|||||
Db 1091 GAGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1150
|||||

QY 493 ATACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 552
|||||
Db 1151 ATTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1210
|||||

QY 553 ATGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 612
|||||
Db 1211 ATGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1270
|||||

QY 613 GTCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 672
|||||
Db 1271 GTCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1330
|||||

QY 673 GAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 732
|||||
Db 1331 GAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1390
|||||

QY 733 GGTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 792
|||||
Db 1391 GGTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1450
|||||

QY 793 GATGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 852
|||||
Db 1451 GATGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1510
|||||

QY 853 CCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 912
|||||
Db 1511 CCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1570
|||||

QY 913 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 972
|||||
Db 1571 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1630
|||||

QY 973 GTCGATATCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1032
|||||
Db 1631 GTCGATATCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1690
|||||

QY 1033 CACAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1092
|||||
Db 1691 CACAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1750
|||||

QY 1093 CAGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1152
|||||
Db 1751 CAGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1810
|||||

QY 1153 ACCAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1212
|||||
Db 1811 ACCAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1870
|||||

QY 1213 TTTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1272
|||||
Db 1871 TTTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1930
|||||

QY 1273 ACTACGATATGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1332
|||||
Db 1931 ACTACGATATGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1990
|||||

QY 1333 AAATGATGTTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1392
|||||
Db 1991 AAATGATGTTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2050
|||||

QY 1393 CGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1452
|||||
Db 2051 CGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2110
|||||

QY 1453 TATAACATATCCA 1465
```

|||||
Db 2111 AATAGCGGAGCAA 2123

RESULT 10
ATTUFA
LOCUS
DEFINITION A.thaliana tufa gene for elongation factor Tu.
ACCESSION X52256
VERSION X52256.1 GI:22564
KEYWORDS elongation factor Tu; tufa gene.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 2272)
Baldauf, S.L.
Direct Submission
Submitted (21-MAR-1990) Baldauf S.L., Indiana University,
Department of Biology, Jordan Hall, Bloomington, Indiana 47405, U S
A
2 (bases 1 to 2272)
Baldauf, S.L. and Palmer, J.D.
Evolutionary transfer of the chloroplast tufa gene to the nucleus
Nature 344 (6263), 262-265 (1990)
90190846
2314461
PUBMED
COMMENT Data kindly reviewed (08-OCT-1990) by Baldauf S.L.
FEATURES
Location/Qualifiers
1..2272
/organism="Arabidopsis thaliana"
/variety="Lausberg erecta"
/db_xref="taxon:3702"
/clone_lib="Lambda"
415..1845
/codon_start=1
/product="elongation factor Tu precursor"
/protein_id="CAA36498.1"
/db_xref="GI:22565"
/translation="MAISAPAAATSSSSRIICSVSSPSLCPAISTSGKUKTLTSSS
FLPSYLTTSASQSTTRRAARGKFKPKPHVNIIGHVDHGTTLTAAALMA
LASIGSSVAKKYQIDEADAAPERAGITINTAYETENRHYAHVDCPGHADYVKNMI
TGAQMDGAILVSGADGPMPTQKEHLLAKQVCPDMVPLNKEDVDVDAELLELV
LEVPRLSSYEFNGDDIPILSGSALLAVETLTENPKVKRGDKWVDKIYELMDVDDY
IPIPQRTPELFLAYEDVESIGRGTATGRVGRGVKVGTVDLGLBETSYTFT
GVEMFQKILDEALAGDNVGLLLRGIQRADIQRMVLAKPSGSIPTPHKFAIIVLKKE
EGGRHSPPFFAGYRPFQYMTDTDTGKVKTKIMNDKDESKMMPGDRVKIVLIVLPPVA
CEQGMFAIREGGKTVGAGVIGTILE"
415..615
/note="putative"
sig_peptide
mat_peptide
616..1842
/product="elongation factor Tu"
BASE COUNT 599 a 467 c 500 g 706 t
ORIGIN

Query Match 40.9%; Score 687.2; DB 8; Length 2272;
Best Local Similarity 72.2%; Pred. No. 4e-83;
Matches 893; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

QY 211 GTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 270
|||||
Db 610 GTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 669
|||||
QY 271 GGCCATGTCGACCGAGAAAGACCACTCTACCGCGCGGCGGCGGCGGCGGCGGCGG 330
|||||
Db 670 GGTCTGTTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 729
|||||
QY 331 GTCCGTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 390
|||||
Db 730 ATTGGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 789
|||||

[illegible]


```

QY 1291 AAGGATGAGGAGCGGATGTCATGCGTGGTGACCGTATCAAAATGATTGTTACGCTC 1350
      ||||| ||| ||||| ||||| ||||| ||| ||| ||||| |||||
Db 9909 AAAGATGAAGAGTCAAGATGTTATGCGCGTGATCGAGTGAAGATTGTTGAGCCTT 9968
      ||||| ||| ||||| ||||| ||||| ||| ||| ||||| |||||

QY 1351 ATCCAGCCTGTTGTTGTGACAGGATGAGGTTGTCATCCGTAGGTTGGTGAAGACC 1410
      || ||| ||| ||||| ||| ||| ||||| ||||| ||||| ||| ||| |||||
Db 9969 ATTTGCGCGGTGGTGTGAACAAGGATGAGGTTGTCATCAGAGAAGGAGGAAGACT 10028
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1411 GTTGTGCGGTGTCATCAACAAATCATTCAGTAA 1446
      ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10029 GTTGTGCTGGAGTTATGGGACGATCTCGAATGA 10064
      ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
LOCUS ATCHRV52 198427 bp DNA linear PLN 16-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 52.
ACCESSION AL161552
VERSION AL161552.2 GI:7268789
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana.
          Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 82704)
AUTHORS Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and
          Mayer,K.F.X.
          Unpublished
          2 (bases 67424 to 179368)
          Terry,N., Ardillies,W., Buysshaert,C., Dasseville,R., De Clerck,R.,
          De Keyser,A., Neyt,P., Rouze,P., Van Den Daele,H., Villarroel,R.,
          Gielen,J., Van Montagu,M., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
          Unpublished
          3 (bases 69292 to 69818)
          Volckaert,G., Grymonprez,B., Voet,M., Robben,J., Mewes,H.W.,
          Lemcke,K. and Mayer,K.F.X.
          Unpublished
          4 (bases 178975 to 187911)
          Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
          Unpublished
          5 (bases 187649 to 198427)
          Pohl,T., Weizenegger,T., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
          Unpublished
          6 (bases 1 to 198427)
          EU Arabidopsis sequencing project.
          Direct Submission
          Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
          Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
          lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
          Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
          Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
          E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed
          annotation of this entry and other sequences of chromosomes 3, 4
          and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
          this fragment has an overlap with ATCHRV51 at the 5' end and an
          overlap with ATCHRV53 at the 3' end.
FEATURES
          Location/Qualifiers
            1..198427
              /organism="Arabidopsis thaliana"
              /variety="Columbia"
              /db_xref="taxon:3702"
              /chromosome="4"
            7547..8512
              /gene="AT4g19950"
              join(7547..8329,8375..8512)
              /note="similarity to p48 eggshell protein gene,
              Schistosoma mansoni, PATCH:G45844
              contains EST gb:AI994604.1, R30520, AI992766.1, T20423,
              AA712864, H76323"

```

```

/codon_start=1
/product="putative protein"
/protein_id="CAB78995.1"
/db_xref="GI:7268790"
/translation="MDLPEELQFLNKGILRESTISPOYSLKFTYLTITLPIPLSF
ALLAHSLETQPILAQIDTYPDQDSQLOHEWTVLVVFOFCYIIFLPAFSLSTAAVVF
TVASYTGKPSVSTMSAIPLVLRKFITFLVWSLLMLAYNTVFLFPLVTLIVAVDL
QNVLIASFVSVFVFLVHVHTALWHLASVSVLEPIYGLAAMKSYELLGKTL
NACSMVFYLVHVCFGIAGVGVAVVRGGDYGITAFIRIVAGGFVSVFYVYVCKSFHQEI
DKSALHDHLLGGVLGVEYVPLKNIQMFENF"
7547..8329
/gene="AT4g19950"
/number=1
8330..8374
/gene="AT4g19950"
/number=1
8375..8512
/gene="AT4g19950"
/number=2
11377..14567
/gene="AT4g19960"
join(11377..11526,11621..11846,11925..12170,12258..12332,
12383..12643,12687..12887,12986..13288,13407..13929,
14024..14567)
/gene="Ar4g19960"
/note="similarity to putative potassium transporter AtKT2p
& AtKTLp, Arabidopsis thaliana, Patchx:G2384669 &
Patchx:G2384671
contains EST gb:AI997834.1"
/codon_start=1
/product="potassium transporter-like protein"
/protein_id="CAB78996.1"
/db_xref="GI:7268791"
/translation="MAERVEASSPEGENTIEEREVGAMWELEKLDQPMDEANKLN
NMYRKGLSMLMLLSFQSLGIVVDLGTSPLYVYTFDPDSDSDVIGALSLLI
YSLLIPLIKYVFIVCANDNGOGTLAIYSLGRHAKVLIHQHRSDELTTYSRT
VSAEGSPAATKKWLEGEKWRKRALLVVLLGTCMIGDGLITPAISGTGKYNPK
MSGNHLFTLFTSDIVVVAIVLILGLFSMQHYGTDKVGWLFAPVILVFLFICATG
MYNICKYDTSVLKAFSPITYIYLFKRRGRDQVLSLGLLSITAIKCYFLNIAGTEA
LYADIAPFLAIQIAFTFFPCLLAYCGAAYLVIRKHYQDAFYASIPDSVYWP
MFIVATGAAIVGSOATISGTYSIVKQAVAHGCFPRKIVHTSKFPLGQIYCPDINWIL
MLGCIATVASFQKQSOIGNAYGKMTTSKYKNKFSOWTAVVLMVLTLLMVLIMLL
VHCHWILVLIFTLSFVSELSYFSAVIFKIDEGWVPLIIAAILSLVMSVYHATVK
KYEFEMHRSVSMWILGSLGVRVPGIYVTELASGVPHIFSHFITNLPAHVS
VYFVCYLPVTTPPEERFLVIRIGPKTRFRFCVARYGYKDLHKDDDDFENKLLTK
LSSFIRITMEPTSSNSTYSSTYSVNHQDSTVDLIHNNHNNHNNHNNMDFFSSMDY
TVSLTDITVSAESLHNTYFSQDNTVEEETDELEFLTKCESGVVHMGNTVVKART
GSLWPKIAIDYVYVAFKICRANSVILHVPHTLLNVGQVYV"
11377..11526
/gene="AT4g19960"
/number=1
11527..11620
/gene="AT4g19960"
/number=1
11621..11846
/gene="AT4g19960"
/number=2
11847..11924
/gene="AT4g19960"
/number=2
11925..12170
/gene="AT4g19960"
/number=3
12171..12257
/gene="AT4g19960"
/number=3
12258..12332
/gene="AT4g19960"
/number=4
12333..12382
/gene="AT4g19960"
/number=4
12383..12643

```


[illegible]

QY	1291	AAGGATGAGAGCGAAGATGTGCATCGCTCGGTGACCGTATCAAAATGATTGTTGAGCTC	1350
Db	188883	AAAGATCAAGAGTCGAAGATGTTATGCCCGGTGATCGACTGAAGATTGTTGAGCTT	188942
QY	1351	ATCCAGCCTGTTGCTTGTGAGCAGGATATGAGTTTGCTATCCGTGAGGTTGGTAAGACC	1410
Db	188943	ATTGTGCCGTGGCTTGTGAACAAGGATGAGTTTGCTATCAGAGAAGGGAAGACT	189002
QY	1411	GTTGGTGC CGGTGCTCATCAACAAATCATTTAGTAA	1446
Db	189003	GTTGGTGTGGAGTTATTTGGGACGATCCTCGAATGA	189038

Search completed: May 1, 2003, 09:37:37
Job time : 5544 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: May 1, 2003, 15:58:03 ; Search time 480 Seconds

(without alignments)
7886.686 Million cell updates/sec

Title: US-09-810-764A-6

Perfect score: 1681

Sequence: 1 attcccaataatccccacc.....gttaaaaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

Result No.	Score	Query Match %	Length	ID	Description
1	1405.4	83.6	1657	21 AAC46577	Zea mays DNA fragm
2	687.2	40.9	1651	21 AAC37985	Arabidopsis thalia
3	685.6	40.8	1651	21 AAC51055	Arabidopsis thalia
4	528.4	31.4	1194	23 AAS54260	Pseudomonas aerugi
5	527.8	31.4	1194	13 AAQ20215	Sequence of tuf1 g
6	526.2	31.3	1194	13 AAQ20219	Sequence of Srtufr
7	526.2	31.3	1194	13 AAQ20218	Sequence of Srtufr
8	524.6	31.2	1194	13 AAQ20220	Sequence of Srtufr
9	524	31.2	29379	23 AAS59510	Propionibacterium

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

10	521.6	31.0	1194	23 AAS54267	Pseudomonas aerugi
11	518.2	30.8	1194	13 AAQ20221	Sequence of Srtufr
12	508.8	30.3	1185	21 AAAG6894	E. coli proliferat
13	508.8	30.3	1185	22 AAH00707	Escherichia coli n
14	508.8	30.3	1185	23 AAS52628	E. coli DNA for ce
15	505.6	30.1	1230	23 AAS56229	Salmonella typhi D
16	503.6	30.0	1185	23 AAS52697	E. coli DNA for ce
17	498	29.6	1191	22 AAH00872	Bordetella pertuss
18	490	29.1	1194	13 AAQ20216	Sequence of tuf2 g
19	488.4	29.1	4411529	22 AA199682	Mycobacterium tube
20	484.4	28.8	1109	22 AAH01846	Klebsiella oxytoca
21	477.4	28.4	4403765	22 AA199683	N. meningitidis pa
22	472	28.1	65632	21 AAA81502	Neisseria meningit
23	472	28.1	349980	21 AAF21544	N. meningitidis B
24	472	28.1	1437668	21 AAA81490	N. meningitidis pa
25	462	27.5	17782	21 AAA81530	Obesumbacterium pr
26	460.6	27.4	1116	22 AAH01849	Buttiauxella agres
27	455.8	27.1	1116	22 AAH01845	C. glutamicum codin
28	450.2	26.8	349980	22 AAH68525	Corynebacterium gl
29	449.4	26.7	1191	22 AAH00677	Elongation factor
30	446.6	26.6	1351	18 AAT92619	C. glutamicum codin
31	446.4	26.4	1188	22 AAH68402	Plesiomonas shigel
32	443.4	26.4	1108	22 AAH01847	Helicobacter pylor
33	436	25.9	1200	23 AAS53681	Genomic fragment #
34	436	25.9	1448	19 AAV07964	Helicobacter pylor
35	433.4	25.8	49617	22 AAF28541	Proteus mirabilis
36	432	25.7	1129	22 AAH01850	Budivicia aquatica
37	421.8	25.1	1112	22 AAH01533	Haemophilus influe
38	420.4	25.0	1185	23 AAS53335	Haemophilus influe
39	420.4	25.0	1185	23 AAS53335	Haemophilus influe
40	420.4	25.0	1830121	17 AAT42063	Haemophilus influe
41	420.4	25.0	1830121	17 AAT42063	Haemophilus influe
42	418.4	24.9	1170	13 AAQ20217	Sequence of tuf3 g
43	413.8	24.6	1057	24 ABK74018	Bacillus lichenifo
44	405.4	24.1	1107	22 AAH01848	Shewanella putrefa
45	402.8	24.0	1245	24 ABN93077	Staphylococcus epi

ALIGNMENTS

RESULT 1	
AAC46577	
ID AAC46577 standard; DNA; 1657 BP.	
XX AAC46577;	
AC AAC46577;	
XX AAC46577;	
DT 18-OCT-2000 (first entry)	
XX Zea mays DNA fragment SEQ ID NO: 50647.	
DE Zea mays subsp. mays.	
XX Hybridisation assay; genetic mapping; gene expression control;	
KW protein identification; signal transduction pathway; metabolic;	
KW pathway; promoter; termination sequence; corn; ss.	
XX	
OS Zea mays subsp. mays.	
XX	
PN Ep1033405-A2.	
XX	
PD 06-SEP-2000.	
XX	
PF 25-FEB-2000; 2000EP-0301439.	
XX	
PR 25-FEB-1999; 99US-0121825.	
PR 05-MAR-1999; 99US-0123180.	
PR 09-MAR-1999; 99US-0123548.	
PR 23-MAR-1999; 99US-0125788.	
PR 29-MAR-1999; 99US-0126264.	
PR 29-MAR-1999; 99US-0126785.	
PR 01-APR-1999; 99US-0127462.	
PR 06-APR-1999; 99US-0128234.	
PR 08-APR-1999; 99US-0128714.	
PR 16-APR-1999; 99US-0129845.	

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 23-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19375.
xx Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 03-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999: 99US-0151303.
 PR 31-AUG-1999: 99US-0151438.
 PR 01-SEP-1999: 99US-0151930.
 PR 07-SEP-1999: 99US-0152363.
 PR 10-SEP-1999: 99US-0153070.
 PR 13-SEP-1999: 99US-0153758.
 PR 15-SEP-1999: 99US-0154018.
 PR 16-SEP-1999: 99US-0154039.
 PR 20-SEP-1999: 99US-0154779.
 PR 22-SEP-1999: 99US-0155139.
 PR 23-SEP-1999: 99US-0155486.
 PR 24-SEP-1999: 99US-0156559.
 PR 28-SEP-1999: 99US-0156458.
 PR 29-SEP-1999: 99US-0156596.
 PR 04-OCT-1999: 99US-0157117.
 PR 05-OCT-1999: 99US-0157753.
 PR 06-OCT-1999: 99US-0157865.
 PR 07-OCT-1999: 99US-0158029.
 PR 08-OCT-1999: 99US-0158232.
 PR 12-OCT-1999: 99US-0158369.
 PR 13-OCT-1999: 99US-0159293.
 PR 13-OCT-1999: 99US-0159294.
 PR 13-OCT-1999: 99US-0159295.
 PR 14-OCT-1999: 99US-0159329.
 PR 14-OCT-1999: 99US-0159330.
 PR 14-OCT-1999: 99US-0159331.
 PR 14-OCT-1999: 99US-0159637.
 PR 14-OCT-1999: 99US-0159638.
 PR 18-OCT-1999: 99US-0160741.
 PR 21-OCT-1999: 99US-0159584.
 PR 21-OCT-1999: 99US-0160767.
 PR 21-OCT-1999: 99US-0160768.
 PR 21-OCT-1999: 99US-0160770.
 PR 21-OCT-1999: 99US-0160814.
 PR 21-OCT-1999: 99US-0160815.
 PR 22-OCT-1999: 99US-0160980.
 PR 22-OCT-1999: 99US-0160981.
 PR 23-OCT-1999: 99US-0160989.
 PR 25-OCT-1999: 99US-0161404.
 PR 25-OCT-1999: 99US-0161405.
 PR 25-OCT-1999: 99US-0161406.
 PR 26-OCT-1999: 99US-0161359.
 PR 26-OCT-1999: 99US-0161360.
 PR 26-OCT-1999: 99US-0161361.
 PR 28-OCT-1999: 99US-0161920.
 PR 28-OCT-1999: 99US-0161992.
 PR 28-OCT-1999: 99US-0161993.
 PR 29-OCT-1999: 99US-0162142.

Query Match 40.9%; Score 687.2; DB 21; Length 1651;
 Best Local Similarity 72.2%; Pred. No. 9.9e-135;
 Matches 893; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

QY 211 GTGCGCGCGGAGGCGCAAGTTCGAGCGGACCAACACACACGCTCAACATAGGACCAATC 270
 DB 261 GTCCGCGCGCGTGTGAAAGTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 320
 QY 271 GGCCATCTGACCAACGAAAGACCACTCTCCCGCGCGGCTCACCATGCTGCTGCTCC 330
 DB 321 GGTGATCTTGACCATGGAAGAACTACTTTAACCAGAGCTCTAACCAGGCTCTCGCTTC 380
 QY 331 GTCCGCGGAGCGGCTTAAAGATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 390
 DB 381 ATTTGGTTCCAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 440
 QY 391 CGCGGTATACCATCAACACCGGCGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 450
 DB 441 CGTGGTATACATCAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
 QY 451 CAGCTGACTGCCCCGCCCGGCGGAGTATGCTCAAGAAATATGATCAACCGCGCTGCGCAG 510
 DB 501 CACGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560

QY 511 ATGAGCGGTGCCATCTCTGCTATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 570
 DB 561 ATGAGCGGAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 620
 QY 571 CACATCTCTCTGCGCAAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 630
 DB 621 CATATCTCTTGGCTTAAGCAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 680
 QY 631 GACATGCTCGAGCAGGAGGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGCTG 690
 DB 681 GATCAAGTAGATGCTGAGAGCTGCTGAGAGCTGCTGAGAGCTGCTGAGAGCTGCTGCTG 740
 QY 691 AGCAACTAGAGTACGAGCGGAGCAGCTACCAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
 DB 741 TCGCTTATGAATTTAACGGTGATGATATCGGATATCTCTGGTCTGCTCTTTAGCC 800
 QY 751 CTCGAGGCTCTATGCTCAACCTGCTTGAAGCGCGGAGGATGATGCTGCTGCTGCTGCTG 810
 DB 801 GTTGAGACTCTTACTGAGAATCTTAAGGTTAAGAGAGGTGATAACAAATGGGTAGATAAG 860
 QY 811 ATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
 DB 861 ATTTATGAACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920
 QY 871 CTCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 930
 DB 921 TTGCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 980
 QY 931 ACTGCCGCTATAGAGCGTGGCACCTCAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 990
 DB 981 ACAGGCGGTGCTGAGAGAGGTACGCTTAAGGTAGGAGAGCTGCTGCTGCTGCTGCTGCTG 1040
 QY 991 CGGGACACCCGAACTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
 DB 1041 AGGGAGACTAGGAGTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
 QY 1051 GCCATGGCGGAGACAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1110
 DB 1101 GCTTTAGCTGGTGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160
 QY 1111 AGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1170
 DB 1161 AGAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1220
 QY 1171 GTGATGCTTTAAGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1230
 DB 1221 ATCTATGCTTTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1280
 QY 1231 CCACAGTTCTACATCGGCAACTGATGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290
 DB 1281 CCTCAGTTCTACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1340
 QY 1291 AAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1350
 DB 1341 AAGATGAAGAGTCTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1400
 QY 1351 ATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1410
 DB 1401 ATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1460
 QY 1411 GTTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
 DB 1461 GTTGTGCTGAGGATTTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1496

RESULT 3
 AAC51055
 ID AAC51055 standard; DNA; 1651 BP.
 XX
 AC AAC51055;
 XX
 DT 18-OCT-2000 (first entry)
 XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67105.
 XX Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125786.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.

DE	Pseudomonas aeruginosa DNA for cellular proliferation protein #391.	
XX	Antisense; ds; prokaryotic cellular proliferation gene;	
KW	antibiotic; antibacterial; drug design.	
XX		
OS	Pseudomonas aeruginosa.	
XX		
PN	W0200170955-A2.	
XX		
PD	27-SEP-2001.	
XX		
PF	21-MAR-2001; 2001WO-US09180.	
XX		
PR	21-MAR-2000; 2000US-191078P.	
PR	23-MAY-2000; 2000US-206848P.	
PR	26-MAY-2000; 2000US-207727P.	
PR	23-OCT-2000; 2000US-242578P.	
PR	27-NOV-2000; 2000US-253252P.	
PR	22-DEC-2000; 2000US-257931P.	
PR	16-FEB-2001; 2001US-269308P.	
XX		
PA	(ELIT-) ELITRA PHARM INC.	
XX		
PI	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;	
PI	Yamamoto RT, Xu HH;	
XX		
DR	WPI; 2001-611495/70.	
XX	P-PSDB; AA036401.	
DR		
XX		
PT	New polynucleotides for the identification and development of	
PT	antibiotics, comprise sequences of antisense nucleic acids -	
XX		
PS	Claim 27; Seq ID No 7897; 511pp; English.	
XX		
CC	The invention relates to antisense inhibitors of genes essential to	
CC	prokaryotic cellular proliferation, their use in identifying the	
CC	genes, their use in the discovery of novel antibiotics, the essential	
CC	genes themselves and the encoded proteins. The prokaryotes used are	
CC	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella	
CC	pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The	
CC	invention is also useful for the identification of potential new targets	
CC	for antibiotic development. The antisense nucleic acids can also be used	
CC	to identify proteins used in proliferation, to express these proteins,	
CC	and to obtain antibodies capable of binding to the expressed proteins.	
CC	The proteins can be used to screen compounds in rational drug discovery	
CC	programmes. The antisense nucleic acid sequence is also useful to screen	
CC	for homologous nucleic acids which are required for cell proliferation in	
CC	a wide variety of organisms. The present sequence encodes an	
CC	essential prokaryotic cellular proliferation protein.	
CC	Note: The sequence data for this patent did not form part	
CC	of the printed specification, but was obtained in electronic	
CC	format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 1194 BP; 259 A; 361 C; 351 G; 223 T; 0 other;	
	Query Match 31.4%; Score 528.4; DB 23; Length 1194;	
	Best Local Similarity 66.1%; Pred. No. 1.9e-101;	
	Matches 813; Conservative 0; Mismatches 381; Indels 36; Gaps 2	
QY	217 GCGGGGAGGGGCAAGTTCGAGGCGCACCAACACACAGCTCAACATAGCACCATCGGCCAT 276	
Db	1 GTGGCTAAGGAAAAATTCGAACGTAAACAACCGCAGGTCAACGTGCGCACCATCGGTAC 60	
QY	277 GTGACACGGAAGAACCACTCTCACCGCGGCTCACCATGGTGTGCTCGCTCCGTCGGT 336	
Db	61 GTTGACCATGGCAAGAACCACTGTACCGCTGCATCGACTGACCAAGGTCTCTCCGATACGTCG 120	
QY	337 GGCACGCGCCTAAGAAGTACGACGAGATCGACGCGCCGCCGAGGAGCGCGCCCGGT 396	
Db	121 GGTGTTCGCTGCTGCTTTCATCGATCGAACACGCGCGGAAGAAAAAGGCCGCGT 180	
QY	397 ATCACCATCAACACCGCCACCGCTCGAGTACGAGACCGGACCCGCCCACTACGACACGTC 456	

AAQ20215
 ID AAQ20215 standard; DNA; 1194 BP.
 AC AAQ20215;
 XX
 DT 15-APR-1992 (first entry)
 DE Sequence of tufI gene encoding translation elongation factor TufI.
 XX
 KW Elfamycin resistant actinomycetes; antibiotic resistant;
 XX elongation factor; ss.
 XX Streptomyces ramocissimus.
 OS
 FH Key Location/Qualifiers
 FT CDS 4..1194
 FT /*tag= a
 XX
 PN EP466251-A.
 XX
 PD 15-JAN-1992.
 XX
 PF 02-JUL-1991; 91EP-0201702.
 XX
 PR 02-JUL-1991; 91EP-0201702.
 PR 10-JUL-1990; 90EP-0201851.
 XX
 PA (KONN) GIST-BROCADES NV.
 XX
 PI Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 PI Woudt LP;
 DR WPI; 1992-017874/03.
 DR P-PSDB; AAR20242.
 XX
 PT New protein conferring resistance to elfamycin - used to
 PT transform streptomycetes to resistant pheno-type
 XX
 PS Example; Fig 1 and Pages 13-15; 35pp; English.
 CC
 CC Substitution of residue 378 of the elongation factor (EF-Tu) with a
 CC valine, threonine, proline or phenylalanine results in an elfamycin
 CC resistant protein (EF-TuR). The advantage of this change is that
 CC the limiting factor for the prodn. of elfamycin by actinomycetes is
 CC removed by mutating the gene tuf into tufR encoding a protein.
 CC resistant to the elfamycin, pref. mocimycin (Kirromycin). The
 CC inventors claim EF-TuR and the genes (tufR) encoding it.
 XX
 SQ Sequence 1194 BP; 238 A; 405 C; 374 G; 177 T; 0 other;
 Query Match 31.4%; Score 527.8; DB 13; Length 1194;
 Best Local Similarity 67.2%; Pred. No. 2.6e-101;
 Matches 833; Conservative 0; Mismatches 352; Indels 54; Gaps 4;
 QY 217 GCGGCGAGGGGCAAGTTCGAGCGCACCAACACACAGTCACATAGGACCATCGGCCAT 276
 DB 1 GTGGCGAAGCGCAAGTTCGAGCGGCACTAAGCGCACGTCACATAGGACCATCGGTCAC 60
 QY 277 GTCGACCGAGGAAGACCACTCTACCGCGCGGCTCACCATTGGTCTCTCTCTCTCTCT 330
 DB 61 ATCGACCGAGTAAAGACGACCTCTACGCGCGGCTTACCAAGGTGCTGACGACGCGTAC 120
 QY 331 GTCGGTGGCAGCGCGCTTAAGAGTACGACGAGATCGACGCGCGCGCGGAGGAGCGCGCC 390
 DB 121 CCGGACCTGACGAGCGCCACCCGCTTCGACACATCGAAGGCTCCTGAGGAGCGTCAG 180
 QY 391 CGCGGTATCACCATCAACACCGCCCGCTCGAGTACGAGCCGAGACCGCGCCACTACGGA 450
 DB 181 CGCGGTATCACCATCTCCATCGCGCACGTCGAGTACGAGACCGGCGGCTCAGTACGCC 240
 QY 451 CAGCTCGACTGCCCCCGGCGCGGCTATGTCAAGATATGATCATCCGCGCGCTGGGAG 510
 DB 241 CAGCTCGACTGCCCCCGGCTACCGGGACTACATCAAGAACATGATCATCGGCTGCGCGCAG 300
 QY 511 ATGGAGGTCGCCATCTCTGTCGTATCCGTCGCGAGGGGCCCATCGCGCAGACCAAGAG 570
 DB 301 ATGGAGGTCGCCATCTCTGTCGTATCCGTCGCGACCGCCGATCGCGCAGACCAAGAG 360
 QY 571 CACATCTCTCGCCCAAGCAAGTCGGTTCCTCAAGATCGTGTCTTCTTCTCAACAAGAG 630
 DB 361 CAGCTCTCTCGCCCGCCAGGTCGGCTTCGTATCGTGTCTGCTCGCCCTCAACAAGGCC 420
 QY 631 GACATGTCGACGACGAGGAGTGTCTGAGCTGCTGAGCTGCTGAGTCTCGCGAGCTGCTC 690
 DB 421 GACATGTCGACGACGAGGAGATCATGAGCTGCTGAGTCTGAGTCTGAGTCTGAGTCTC 480
 QY 691 AGCAACTAGGAGTACGACGCGGACGACGACGACGACGACGACGACGACGACGACGACG 750
 DB 481 TCCGAGTACGAGTTCCCGGCGGACGACGACGACGACGACGACGACGACGACGACGACG 540
 QY 751 CTCGAGGCTCTCATGTGTCACACCTGCTTGAAGCGCGGCGGACGATGAGTGGTTCGACTAC 810
 DB 541 CTGGAG-----GGCGACGCTCAGTGGACGACGCTC 570
 QY 811 ATCTTCTGTTGTTGATATAAGTGGATTCTTATATTTCCAGTCCCGCAGAGGACGACTGAC 870
 DB 571 GTCTCGACCTGATGAAGCGCGCTCGACGAGTCCATCCCGGAGCGGAGCGGACGCTCGAC 630
 QY 871 CTCCCGTCTGCTGCTGTTGAAGATGCTTCTTCCATCACCCTGCTGCTGCTGCTGCTGCT 930
 DB 631 AAGCGGTTCTCATGCGGATCGGAGGAGGCTTTCAGCATCACCCTGCTGCTGCTGCTGCTG 690
 QY 931 ACTGGCGGTATAGAGCGTGGCACCCTGCAAGATTGGTGACACAGTATGCTCGGAATC 990
 DB 691 ACCGGCGTATCGAGCGTGTGCTCTGAGGTCAAGGACGACGACGACGACGACGACGACG 750
 QY 991 CGG---GACACCCGGAACCTGACGCTCAGTGTGTTGAGATGTTCCAGAGACCATGGAT 1047
 DB 751 AAGACCGAGAGAACACACCGGTCACCGGATCGAGATGTTCCGCAAGTGTCTCGAC 810
 QY 1048 GATGCCATGGCGGAGACAATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107
 DB 811 GAGGCGCAGCGCGGTGAGAACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
 QY 1108 GAAAGAGGATGTTGCTGGCAAGCCCTGCTATACACCGGACACACCAAGTTGAGGCT 1167
 DB 871 GAGCGCGCGGAGGTCATCAAGCGCGGCTCGGTCACCGGACACCGGAGTTCGAGGCG 930
 QY 1168 GTTGTGTATGCTTAAGAGAGAGGAGGTCGCGGACACCTACCTTTCTTCCCTGGTAC 1227
 DB 931 CAGGCTACATCTCTTCCAGGAGGAGGTCGCGGCGGACACGCGGTTCTTCAACAACATC 990
 QY 1228 CGCCACAGTCTACATGCGGACAACTGATGTGACAGGAGTGTGACTACGATTGAT 1287
 DB 991 CGCCCGAGTGTCTTCTTCCGTCACCGGAGCTGACCGGCTGTTGTG----- 1035
 QY 1288 GACAAGATGAGGAGGAGGAGATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1347
 DB 1036 CACCTCCCGGAGGAGCAGGAGATGTCATCGCGGCGGACACACCGAGATGCGGCTCGAG 1095
 QY 1348 CTATCTCAGCTGTTGCTGTGAGCAGGATGATGAGTTTGTCTATCTGAGGAGGTTGAG 1407
 DB 1096 CTGATCCAGCGGTCGCCATGAGGAGGAGGCTGAAGTTCGCGCATCCGCTGAGGTTGCG 1155
 QY 1408 ACCGTTGTCGCGGTGTCATCAACAATAATCTTCTGAGTAA 1446
 DB 1156 ACCGTCGCGCGCGGCGGAGGTCACCAAGATCGTCAAGTAA 1194
 RESULT 6
 AAQ20219
 ID AAQ20219 standard; DNA; 1194 BP.
 AC AAQ20219;
 XX
 XX 15-APR-1992 (first entry)

XX DE Sequence of SrtufR1 gene encoding elfamycin-resistant elingation
 XX factor EF-TuR Thr 378.
 KW Elfamycin resistant actinomycetes; antibiotic resistant;
 KW elongation factor; ss.
 XX OS Streptomyces ramocissimus.
 XX FH Key Location/Qualifiers
 FT CDS 4..1194
 FT /*tag= a
 XX PN EP466251-A.
 XX PD 15-JAN-1992.
 XX PF 02-JUL-1991; 91EP-0201702.
 XX PR 02-JUL-1991; 91EP-0201702.
 XX PR 10-JUL-1990; 90EP-0201851.
 XX PA (KONN) GIST-BROCADES NV.
 XX PI Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
 XX PI Woudt LP;
 XX DR WPI: 1992-017874/03.
 XX DR P-PSDB; AAR20246.
 XX PT New protein conferring resistance to elfamycin - used to
 PT transform streptomycetes to resistant pheno-type
 XX PS Claim 8; Pages 13-15; 35pp; English.
 XX CC Substitution of residue 378 of the elongation factor (EF-Tu) with a
 CC valine, threonine, proline or phenylalanine results in an elfamycin
 CC resistant protein (EF-TuR). The advantage of this change is that
 CC the limiting factor for the prodn. of elfamycin by actinomycetes is
 CC removed by mutating the gene tuf into tufR encoding a protein
 CC resistant to an elfamycin, pref. mociimycin (Kirromycin). The
 CC inventors claim EF-TuR and the genes (tufR) encoding it.
 XX SQ Sequence 1194 BP; 239 A; 405 C; 373 G; 177 T; 0 other;

Query Match 31.3%; Score 526.2; DB 13; Length 1194;
 Best Local Similarity 67.2%; Pred. No. 5.6e-101;
 Matches 832; Conservative 0; Mismatches 353; Indels 54; Gaps 4;

QY 217 GCGGCGAGGGCAAGTTCGAGCGCCACCAACACACGTCACATAGGACCATCGGCCAT 276
 DB 1 GTGGCGAAGGCGAAGTTTCGAGCGGACTAAGCGGCACGTCACATGGGCACCATCGGTAC 60

QY 277 GTCGACACGCGAAGAACCACTCTACCGCGCGCTCACCATTGGTGCT-----CGCCTCC 330
 DB 61 ATCGACACCGGTAAACAGACCCCTCAGCGCGCCATTACCAAGGTGCTGCACAGCGGTAC 120

QY 331 GTCGTGGCAGCGCGCTTAAGAAGTACGACGAGATCGACGCGCGCCCGGAGAGCGGCC 390
 DB 121 CGGACCTGAACGAGGCGCACCCCGTTTCGACACATCGACAAAGGCTCCTGAGGAGCGTCAG 180

QY 391 CGCGGTATCACCATCAACACCGCCACCGTCGAGTACGAGACCGCCGCTACGCA 450
 DB 181 CGCGGTATCACCATCTCCATCCGACGCTCGAGTACCGACCGAGCGCGGTACGAGCC 240

QY 451 CACGTGCACTGCGCGCGCCAGCGGCTATGTCAAGATATGATCACCAGCGGTGCGCAG 510
 DB 241 CACGTGCACTGCGCGGTGTCACGCGGACTATCATCAAGACATGATCAGGTTGCGCGCAG 300

QY 511 ATGGACGTGCGCATCTCGTATCGGTATCGGTGCGGACGCGGCCATGCCGACGACCAAGAG 570
 DB 301 ATGGACGCGCGCATCTCTGTGTGTCGCGGCCACCGACGCGCGCATGCCGACGACCAAGAG 360

QY 571 CACATCTCTCTCGCCCAAGCAAGTCGGTGTTCACAGATCGTTGTCTTCTCTCAACAAGAG 630
 DB 361 CACGTGCTCTCGCCCGCAGGTCGGCGTTCGTATACATCGTGGTCTCGTCAACAAGGCC 420

QY 631 GACATGGTGGACGACGAGGAGCTGTCGAGCTCGTCGAGCTCGGAGTCCGCGAGCTGTC 690
 DB 421 GACATGGTGGACGACGAGGAGATCATGGAGCTCGTTGAGCTCGAGGTCGAGCTCCTC 480

QY 691 AGCAACTACGAGTACGACGCGGACGACGTCAGCAATCGCTCGCTGGCTCCGCCCTCAAGCG 750
 DB 481 TCCGAGTACGAGTTCGCGGCGGACGACCTGCGGTCTCGCGTCTCGCGCTGAAGCG 540

QY 751 CTCGAGGCTCTCATGGTCAACCCCTGCTTGAAGCGCGGACGATGAGTGGGTGCGACTAC 810
 DB 541 CTGGAG-----GGCGACGCTCAGTGCACGAGTCC 570

QY 811 ATCTTCTCGTTGTGATAAAGTGGATTCCTATATTCAGTCCCGCAGAGCGACACTGAC 870
 DB 571 GTCTCGACCTGATGAAGGCGCTGACGAGTCCATCCGCGAGCGGAGCGGACGCTCGAC 630

QY 871 CTCGCGTTCTTCTCGCTGTTGAAGATGCTCTTCCATCACCGCTCGTGGTGTACAGTTGCC 930
 DB 631 AAGCGTTCTCTCATGCGGATCGAGGAGCTCTTACGATCACCGTTCGCGGACGCTGTC 690

QY 931 ACTGGCGTATAGAGCGTGGCACCGTCAAGATTTGGTGTGACAGTCTGATATGTCGGAATC 990
 DB 691 ACCGCGGTATCGAGCGGTGTCTCTGAAGGTCAACGAGACCGTCGACATCATCGGCATC 750

QY 991 CGG---GACACCGGAACTGCACGCTCAGTGTCTTCCAGATGTTCCAGAGACCATCGAT 1047
 DB 751 AAGACCGAGAGACCAACACCGTACCGGATCGAGATGTTCCGCAAGCTGCTCGAC 810

QY 1048 GATGCCATGGCGGAGACAATGTTGGCTGCTGCTCGTGTATGACAGAGGATGACATT 1107
 DB 811 GAGGCGCAGGCGGTGAGAACGTCGGTCTGCTGCTCGCGCATCAAGCGGAGGAGTC 870

QY 1108 GAAAGAGCATGGTGTGGCAAGCCCTGCTCTATCACCGGACACCAAGTTTGAAGCT 1167
 DB 871 GAGCGCGCAGGTCATCATCAAGCCGGGTGCGTCAACCCCGCACACCGAGTTGAGGG 930

QY 1168 GTTCTGTATGTCTTAAGAAGAGAGGTTGGCGCGCACACTCAGCTTTCTTCCCTGGTTAC 1227
 DB 931 CAGGCTTACATCTCTCCAAAGGAGGAGGTTGGCGCCACACGCGGTCTTTCACAACTAC 990

QY 1228 CGCCACAGTTCTACATCGCGACAATGATGTGACAGGAGTGTGACTACGATTATGAAT 1287
 DB 991 CGCGCGCAGTTCTACTTCCGTACCGACGAGTGGCGCCACGCGGTCTTTCACAACTAC 1035

QY 1288 GACAAGGATGAGGAGGCAAGATGTGATCGCTGGTGACCGTATCAAAATGATTGTCAG 1347
 DB 1036 CACCTCCCGGAGGCGACCGAGATGGTATCGCGGGGCAACACCGAGATGCGCGTCGAG 1095

QY 1348 CTCATCAGCCTGTGTTGTGAGCAGGATATCAGGTTTCTATCCGTGAGGTTGTAAG 1407
 DB 1096 CTGATCCAGCCCGTCGCGCATGAGGAGGCGCTCAAGTTCACCATCTCGTGGGTGCGCG 1155

QY 1408 ACCGTGGTCCCGTGTCTACCAAAATCATTTGAGTAA 1446
 DB 1156 ACCGTGGCGCGCGACGAGTCAACCAAGATGCTCAAGTAA 1194

RESULT 7

AAQ20218

ID AAQ20218 standard; DNA; 1194 BP.

XX AAQ20218;

AC AAQ20218;

DT 15-APR-1992 (first entry)

XX Sequence of SrtufR1 gene encoding elfamycin-resistant elingation
 DE factor EF-TuR Val 378.
 XX Elfamycin resistant actinomycetes; antibiotic resistant;

elongation factor; ss.
Streptomyces ramocissimus.
Key Location/Qualifiers
CDS 4..1194
/*tag= a
EP466251-A.
15-JAN-1992.
02-JUL-1991; 91EP-0201702.
02-JUL-1991; 91EP-0201702.
10-JUL-1990; 90EP-0201851.
(KONN) GIST-BROCADES NV.
Luiten RGM, Kerkman R, Bosch L, Vijgenboom E, Heinstra PW;
Woudt LP;
WPI; 1992-017874/03.
P-PSDB; AAR20245.
New protein conferring resistance to elfamycin - used to
transform streptomycetes to resistant pheno-type
Claim 8; Pages 13-15; 35pp; English.
Substitution of residue 378 of the elongation factor (EF-Tu) with a
valine, threonine, proline or phenylalanine results in an elfamycin
resistant protein (EF-TuR). The advantage of this change is that
the limiting factor for the prodn of elfamycin by actinomycetes is
removed by mutating the gene tuf into tufR encoding a protein
resistant to an elfamycin, pref. mocimycin (Kirmomycin). The
inventors claim EF-TuR and the genes (tufR) encoding it.
SQ Sequence 1194 BP; 238 A; 404 C; 374 G; 178 T; 0 other;
Query Match 31.3%; Score 526.2; DB 13; Length 1194;
Best Local Similarity 67.2%; Pred. No. 5.6e-101; Indels 54; Gaps 4;
Matches 832; Conservative 0; Mismatches 353;
QY 217 GCGCGAGGCGGCAAGTTCGAGCGGCAACACACAGTCAACATAGGCGCCATCGGCCAT 276
DB 1 GTGCGAAGCGGAATTCGAGCGGCACTAAGCCGCGTCAACATGGCGCCATCGGTAC 60
QY 277 GTGACACAGGAAGACCACTCTACCGCGCGCTCACATGGTGTCT-----CGCTCC 330
DB 61 ATGACACAGGTAAGACGACCTTCAGGCGCGCATTTACCAAGGTGCTGCACGCGGTAC 120
QY 331 GTGCGTGGCGCGCTTAGAAGTACGAGGATCGAGCGCGCCCGCCCGGAGCGCGCC 390
DB 121 CGGACCTGACGAGCGGCGCCCGCTTCGAGCAACATCGAAGGCTCTTGAGGAGCGTCAG 180
QY 391 CGGCGTATACCAACACCGCCACCGTTCGAGTACGAGACCGCGCCACTACGCA 450
DB 181 CGGCGTATACCACTTCCTCCGCGCAGCTCGAGTACGAGACCGCGGCTCAGTACGCC 240
QY 451 CAGCTGCACTGCCCCGCCCGCGGCACTATGTCGAAGATATGATACCGCGCGCTCGCGAG 510
DB 241 CAGCTGCACTGCCCCGGGTACCGCGGCACTATACCAAGAACATGATCAGCGGTGCGCGCAG 300
QY 511 ATGACCGGTGCCATCTCTCGTATCGGTCGCGAGCGGCCCATGCGCGAGACCAAGAG 570
DB 301 ATGACCGCGCCATCTCTGTTGCGCGCCACCGAGCGCCGATGCGCGAGACCAAGAG 360
QY 571 CACATCTCTCTCGCGAAGCAAGTTCGTTCCCAAGATCTGTTCTCTCTCAACAAGAG 630
DB 361 CAGCTGCTCTGCGCCCGCGAGTTCGCGGTTCCTGATCTGCTGCTGCTGCTGCTGCTGCT 420
QY 631 GACATGTCGACGACGAGGAGTCTCGAGCTCGTCTGAGTCTGAGTCTGCGAGTCTGCTGCT 690

Db 421 GACATGTTGGACGACGAGGATCATGGAGCTCGTTCAGCTCGAGGTCCGTGAGCTCTC 480
QY 691 AGCACTACGAGTACGAGCGGACGACCTACCAATCTCGTTCGCTCGGCTCGGCTCAAGGCG 750
Db 481 TCCGAGTACGAGTTCGCGGCGGAGACCTGCGGTGCTGCGGTGCTCGGCTCGGCTCAAGGCG 540
QY 751 CTCGAGGCTCTCATGGTCAACGCTTGAAGCGCGGACGATGATGATGATGATGATGATGATGAT 810
Db 541 CTGGAG-----GGCGAGCTCAGTGAGCAGCTCC 570
QY 811 ATCTTCTCGTTCGTTGATTAAGTGGATTCCTATATTCAGTCCCGCAGACGACACTGAC 870
Db 571 GTCTCTCGACCTGATGAAGCGGCTCGACGAGTCCATCCCGGAGCGCGGAGCGAGCTCGAC 630
QY 871 CTCGCGTTCCTGCTGCTGTTGAAGATGCTTCTTCATCACCGGTTCGTTGATGATGATGATGATGATGAT 930
Db 631 AAGCGGCTCTCATGCTGATGAGGACGCTTTCAGATCACCGGTTCGCGGCGGACGCTGCTC 690
QY 931 ACTGCGCTATAGAGCGTGGCACCCTCAAGATTCGTTGATGATGATGATGATGATGATGATGATGAT 990
Db 691 ACCGCGCTGATCGAGCGTGGTCTCTGAAGGTCAACGAGACCGTTCGACATCATCGGCATC 750
QY 991 CGG---GACACCGGAACTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1047
Db 751 AAGACCGGAGAGACCAACCGGCTCACCGGCTACCGGCTACCGGCTACCGGCTACCGGCTACCGGCTACCGGCT 810
QY 1048 GATGCCATGCGCGGAGACAATGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107
Db 811 GAGGCGCGGCGGCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
QY 1108 GAAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1167
Db 871 GAGCGCGCGGAGTCTATCATCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
QY 1168 GTTGTGATGCTTAAAGAGAGAGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
Db 931 CAGGCGCTATCTCTTCAAGAGAGAGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 990
QY 1228 CCGCCACAGTTCATCATCGGCAACTGATGTGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1287
Db 991 CGCCCGCGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1035
QY 1288 GACAAGGATGAGAGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1347
Db 1036 CACCTCCCGGAGGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1095
QY 1348 CTCATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1407
Db 1096 CTGATCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1155
QY 1408 ACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
Db 1156 ACCGTCGCGCGCGGCGGAGTCAACGATCGTCAAGTAA 1194
RESULT 8
AAQ20220
ID AAQ20220 standard; DNA; 1194 BP.
XX AAQ20220;
AC AAQ20220;
DT 15-APR-1992 (first entry)
XX
DE Sequence of 'StufR1' gene encoding elfamycin-resistant elongation
DE factor EF-TuR Pro 378.
XX
KW Elfamycin resistant actinomycetes; antibiotic resistant;
XX elongation factor; ss.
OS Streptomyces ramocissimus.
XX
FH Key Location/Qualifiers

RESULT 15

REFUGEE ID
AAS56279

AAS36229
ID AAS56229 standard: DNA: 1230 BP.

XX
XX
XXXXXXX

100

AC AAS56229;

XX
XX
DT 13-FEB-2002 (first entry)
XX
DE Salmonella typhi DNA for cellular proliferation protein #262.

[illegible]

KW Antisense; ds; prokaryotic cellular pro

KW antibiotic; antiba

XX

OS Salmonella t:

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

2 XX
3 XX
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

21-MAR-2001: 2001WO-US09180.

XX
ZT MAY 2001, 2001MC 0305180.

21-MAR-2000. 2000US-191078P

21-MAR-2000; 2000US-19107
23-MAY-2000; 2000US-20684

PK 23-MAY-2000; 2000US-206646P.
PR 26-MAY-2000: 2000US-207727P

PR 26-MAY-2000; 2000US-
PR 23-OCT-2000; 2000US-

PR 23-OCT-2000; 2000US-
PR 27-NOV-2000; 2000US-

PR 27=NOV=2000; 2000

PR 22-DEC-2000; 2000S-25793IP.
DB 16-FEB-2001; 2001NC-2502000

PR 16-FEB-2001; 200105-269308P.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

PA	(ELIT-) ELITRA PHARM INC.	
XX		
PI	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;	
PI	Yamamoto RT, Xu HH;	
XX	WPI; 2001-611495/70.	
DR	P-PSDB; AAU38370.	
DR		
XX		
XX	New polynucleotides for the identification and development of	
PT	antibiotics, comprise sequences of antisense nucleic acids -	
PT		
XX	Claim 27; Seq ID No 9866; 511pp; English.	
XX		
PS		

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acid can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published.pct.sequences](http://wipo.int/pub/published.pct.sequences).

XX	seq = refname / fseq / parname - pos - sequence
SQ	Sequence 1230 BP; 301 A; 323 C; 339 G; 266 T; 1 other;
	Query Match 30.1%; Score 505.6; DB 23; Length 1230;
	Best Local Similarity 65.1%; Pred. No. 1.2e-96;
	Matches 801; Conservative 0; Mismatches 385; Indels 45; Gaps 2;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Qv 216 CGCGGCGAGGGGCAAGTTCGAGCGCACCAACCAACGTCACACATAGGCACCATCGGCCA 275

QY 1356 GCCTGTGCTTGTGACGAGGATATGAGTTTGGCTATCCGTGAGGTTGGTAAGACCGTTGG 1415
 Db 1140 CCGATCGCAATGGACGCGTCTGCGTTTCGCAATCCGTGAAGCGCGCGTACCGTTGG 1199
 QY 1416 TGCCGGTGTATCAACAAAAATCATTGAGTAA 1446
 Db 1200 CGCGGCGTTGTGTAAAGTTCTGGGCTAA 1230

Search completed: May 1, 2003, 23:43:47
 Job time : 505 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 07:06:31 ; Search time 2893 Seconds
(without alignments)
9410.513 Million cell updates/sec

Title: US-09-810-764A-6
Perfect score: 1681
Sequence: 1 attcccaataatccccacc.....gttaaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:

1: em_estba:*
2: em_estin:*
3: em_estin:*
4: em_estin:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1054	62.7	1193	11	AY105509	AY105509 Zea mays
2	569.8	33.9	831	12	BE705560	BE705560 Sc01_03a0
3	563.4	33.5	624	9	AI665388	AI665388 605010F02
4	537.6	32.0	681	14	BQ245968	BQ245968 TaB15017G
5	523.4	31.1	736	10	AV943931	AV943931 AV943931
6	521.2	31.0	761	12	BE705674	BE705674 Sc01_11c1

C	7	520	30.9	728	13	BJ308303	BJ308303
C	8	515.4	30.7	900	12	BG301275	BG301275 HVSMED002
C	9	509.6	30.3	685	13	BJ263237	BJ263237
C	10	507.4	30.2	658	13	BJ302467	BJ302467
	11	501.4	29.8	678	13	BI959687	BI959687 HVSMED002
C	12	500.6	29.8	644	10	AV938814	AV938814 HVSMED002
C	13	498.4	29.6	670	13	BJ255324	BJ255324
C	14	497.6	29.6	632	12	BG908358	BG908358 TaLr1167E
C	15	493.2	29.3	558	13	BI992656	BI992656 1020065H0
C	16	493	29.3	682	13	BJ320206	BJ320206
C	17	491.8	29.3	623	11	AY107645	AY107645 Zea mays
C	18	488.6	29.1	574	10	AM672043	AM672043 LG1_334_D
C	19	487.2	29.0	636	13	BJ217502	BJ217502
C	20	477.8	28.4	606	13	BJ257818	BJ257818
C	21	477.2	28.4	826	12	BF262018	BF262018 HV_CEA000
C	22	475.4	28.3	624	13	BJ249074	BJ249074
C	23	472.6	28.1	650	9	AL507013	AL507013
C	24	471.8	28.1	608	10	AV941424	AV941424
C	25	471.2	28.0	907	10	BE216372	BE216372 HV_CEB001
C	26	469.4	27.9	514	12	BG410506	BG410506 947048F12
C	27	462.8	27.5	639	13	BI118749	BI118749 EST136 D1
C	28	461	27.4	574	13	BJ208789	BJ208789
C	29	458.2	27.3	648	13	BI946293	BI946293 sr70q08.Y
C	30	455.4	27.1	639	14	BQ789395	BQ789395 WHEA160.G
C	31	454.8	27.1	636	9	AJ431953	AJ431953
C	32	447	26.6	637	9	AL500399	AL500399
C	33	441.6	26.3	601	13	BJ229928	BJ229928
C	34	439.8	26.2	532	10	AV922915	AV922915
C	35	434.6	25.9	595	10	AV946486	AV946486
C	36	432.2	25.7	473	12	BG410505	BG410505 947048F12
C	37	431.8	25.7	534	12	BF291714	BF291714 WHE2205_B
C	38	430.2	25.6	467	12	BG349916	BG349916
C	39	429.6	25.6	549	9	AL822168	AL822168
C	40	429.6	25.6	551	12	BG263489	BG263489 WHE2342_F
C	41	423	25.2	511	10	BE470937	BE470937 WHE0282.H
C	42	422.6	25.1	582	10	AV945813	AV945813
C	43	421	25.0	580	12	BG906350	BG906350 TaLr1148G
C	44	420.6	25.0	1815	11	AY104241	AY104241 Zea mays
C	45	418.2	24.9	712	13	BI941934	BI941934 se08g11.Y

ALIGNMENTS

RESULT 1	AY105509	AY105509	1193 bp	mRNA	linear	HTC 25-MAY-2002
LOCUS	Zea mays	PC0100031	mrna	sequence.		
DEFINITION	Zea mays	PC0100031	mrna	sequence.		
ACCESSION	AY105509					
VERSION	AY105509.1	GI:21208587				
KEYWORDS	HTC					
SOURCE	Zea mays					
ORGANISM	Zea mays					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC					
AUTHORS	Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.					
TITLE	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes					
JOURNAL	Unpublished (2002)					
REFERENCE	2 (bases 1 to 1193)					
AUTHORS	Coe, E.C.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA					
FEATURES	Location/Qualifiers					
source	1..1193					
	/organism="Zea mays"					
	/db_xref="MaizeDB:635741"					
	/db_xref="taxon:4577"					
	/clone="PC0100031"					

```

/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT      283 a   290 c   328 g   291 t   1 others
ORIGIN

Query Match      62.7%; Score 1054; DB 11; Length 1193;
Best Local Similarity 93.4%; Pred. No. 4.4e-200;
Matches 1112; Conservative 0; Mismatches 76; Indels 3; Gaps 1;

QY 491 TGATCACCGCGCTGCGCAGATGAGCGGTGCCATCTCTGCTATCCGCTGCGGAGCGC 550
DB 1 TGATCACCGCGCTGCGCAGATGAGCGGTGCCATCTCTGCTATCCGCTGCGGAGCGC 60
QY 551 CCATGCGCGAGACCAAGAGACATCTCTCTCGCCCAAGCAAGTCGTTCCTCAAGATCG 610
DB 61 CCATGCGCGAGACCAAGAGACATCTCTCTCGCCCAAGCAAGTCGTTCCTCAAGATCG 120
QY 611 TTGCTCTTCTCAACAAGAGGACATGCTCGACGAGGAGAGCTGCTCGAGCTGCTGAGC 670
DB 121 TTGCTCTTCTCAACAAGAGGACATGCTCGACGAGGAGAGCTGCTCGAGCTGCTGAGC 180
QY 671 TCGAGTCCGCGAGCTGCTCAGCACTAGAGTACGAGGAGGAGGAGGAGTACCAATCGTCG 730
DB 181 TCGAGTCCGCGAGCTGCTCAGCACTAGAGTACGAGGAGGAGGAGGAGTACCAATCGTCG 240
QY 731 CTGGCTCCCGCTCAAGGCGCTCGAGGCTCTCATGCTCAAGCTGCTGAGGCGGCGG 790
DB 241 CTGGCTCCCGCTCAAGGCGCTCGAGGCGCTCTCATGCTCAAGCTGCTGAGGCGGCGG 300
QY 791 ACGATGAGTGGTGACATCTCTCTGCTGTTGATTAAGTGGATTCCTATATTCAG 850
DB 301 ACGATGAGTGGTGACATCTCTCTGCTGTTGATTAAGTGGATTCCTATATTCAG 360
QY 851 TCCCGCAGAGGAGACTGACCTCCGTTCTGCTCGCTGTTGAAGATGCTTCTCCAPCA 910
DB 361 TCCCGCAGAGGAGACTGACCTCCGTTCTGCTCGCTGTTGAAGATGCTTCTCCAPCA 420
QY 911 CCGGTCGTGGTACAGTGTGCACTGCGCTATAGAGCGTGGACGTCAGGATTCGTCACA 970
DB 421 CCGGTCGTGGTACAGTGTGCACTGCGCTATAGAGCGTGGACGTCAGGATTCGTCACA 480
QY 971 CAGTGCATATCTCGGATTCGGGACACCCGGAACCTGACGCTGCTGTTGAGATGT 1030
DB 481 CAGTGCATATCTCGGATTCGGGACACCCGGAACCTGACGCTGCTGTTGAGATGT 540
QY 1031 TCCAGAAGACCATGATGATGCAATGCGCGGAGACATGTTGGGCTGCTGCTCGTGGA 1090
DB 541 TCCAGAAGACCATGATGATGCAATGCGCGGAGACATGTTGGGCTGCTGCTCGTGGA 600
QY 1091 TCCAGAAGGATGACATTTGAAAGAGCATGTTGCTGGCAAGCCCTGCTTATACACCGC 1150
DB 601 TCCAGAAGGATGACATTTGAAAGAGCATGTTGCTGGCAAGCCCTGCTTATACACCGC 660
QY 1151 ACACCAAGTTTGGAGCTGTTGATGCTTAAAGAGGAGAGGTTGGCGGACACTCAC 1210
DB 661 ACACCAAGTTTGGAGCTGTTGATGCTTAAAGAGGAGAGGTTGGCGGACACTCAC 720
QY 1211 CTTTCTCCCTGGTTACCGCCAGTTCATGCTGCGGACACTGATGCTGACAGGAGTG 1270
DB 721 CTTTCTCCCTGGTTACCGCCAGTTCATGCTGCGGACACTGATGCTGACAGGAGTG 780
QY 1271 TGACTACGATTATGAATGACAAGGATGAGGAGGCAAGATGTCATGCTGCTGACCGTA 1330
DB 781 TGACTGATTATGAATGACAAGGATGAGGAGGCAAGATGTCATGCTGCTGACCGTA 840
QY 1331 TCAAAATGATTGTCAGTCTATCAGCCTGTTGCTGAGAGGAGGATGAGTTGCTA 1390
DB 1391 TCAAAATGATTGTCAGTCTATCAGCCTGTTGCTGAGAGGAGGATGAGTTGCTA

```

```

DB 841 TCAGATGGTGTTCAGCTCATCCAGCCTGTGCTGTGACAGGGTATGAGTTGCTA 900
QY 1391 TCCGAGGGTGGTAAGACCGTGTGTCGCCGTGTCTATCAACAAATCATTTGAGTAACTG 1450
DB 901 TCCGAGGGTGGTAAGACTGTGTCGTGTGTCATCAACAAATCATTTGAGTAACTG 960
QY 1451 GATATAACATATCCACCATGAGAAATTTCCCTTGTCTTACTCAAGCGACATGCTCCGAGT 1510
DB 961 GATATATCATATCCACCATGAGAAATTTCCCTTGTCTTACTCAAGCGAAATGCTCTGAGT 1020
QY 1511 TGTATTATGTGTCAGTGTGTTAGGGTGTGCTCATGTGCAATGTAGTATGACACTTTTT 1570
DB 1021 TGTATTATGTGTCAGTGTGTTAGGGTGTGCTCATGTGCAATGTAGTATGACAC---TCT 1077
QY 1571 TTTGTCAAGTGAATTTGTCATAATTTATGACATTCACGACAAAGATTCACATATGCTGTG 1630
DB 1078 TTTGTCAAGTGAATTTGTCATAATTTATGACATTCACGACAAAGATTCACATATGCTGTG 1137
QY 1631 CAACCTCAATTTGGCTAAGAGTGGCCATCTACTGTTTAAAAAATAAAAAA 1681
DB 1138 CAACCTCAATTTGGCTAAGAGTGGCCATCTACTGTTTAAAAAATAAAAAA 1188

RESULT 2
BE705560/c
LOCUS
DEFINITION
  BE705560      831 bp      mRNA      linear      EST 12-SEP-2000
  Sc01_03a09_A Sc01_AAFc_EC0RC_cold_stressed_winter_rye_seedlings
  Secale cereale cDNA clone Sc01_03a09, mRNA sequence.
ACCESSION
  BE705560
VERSION
  BE705560.1 GI:10093825
KEYWORDS
  EST.
SOURCE
  rye.
ORGANISM
  Secale cereale
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
  ; Triticeae; Secale.
  1. (bases 1 to 831)
  ,J.I., Ouellet,T., Piche,C., Couroux,P., De Moors,A., Harris,L.J., Hattori
  ,J.I., Ouellet,T., Robert,L.S., Spratt,D. and Tinker,N.A.
  Expressed Sequence Tags from Cold-Stressed Winter Rye Seedlings
  Unpublished (2000)
  Contact: Singh,J.A.
  Eastern Cereal and Oilseed Research Centre
  Agriculture and Agri-food Canada
  KW Neathy Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
  0C6, Canada
  Tel: (613) 759-1662
  Fax: (613) 759-1701
  Email: singhja@em.agr.ca.

FEATURES
      source
      1. .831
         /organism="Secale cereale"
         /cultivar="Puma (winter rye)"
         /db_xref="taxon:4550"
         /clone_lib="Sc01_03a09"
         /tissue="leaf, crown"
         /dev_stage="seedling three-leaf stage"
         /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI;
         Site_2: Xho I; Sampled three-leaf seedlings treated for
         one week at 20C, 12 hrs light/day. Library made with
         Stratagene UNIZAP XR Kit/GigaPack III Gold Kit. Lambda
         library is amplified, then mass excised in SOLR cells."
BASE COUNT      210 a   243 c   187 g   190 t   1 others
ORIGIN

Query Match      33.9%; Score 569.8; DB 12; Length 831;
Best Local Similarity 83.9%; Pred. No. 1.4e-103;
Matches 692; Conservative 1; Mismatches 123; Indels 9; Gaps 4;

QY 750 GCTCGAGGCTCTCATGGTCAACCCCTGCTTGAAGCGCGGACGATGAGT-GGGTCGACT 808
DB 750 GCTCGAGGCTCTCATGGTCAACCCCTGCTTGAAGCGCGGACGATGAGT-GGGTCGACT

```


RESULT 4
 BQ245968
 LOCUS
 DEFINITION
 BQ245968
 TAE15017G12R TAE15 Triticum aestivum CDNA clone TAE15017G12R, mRNA
 sequence.
 BQ245968
 VERSION
 BQ245968.1 GI:2041844
 EST.
 SOURCE
 bread wheat.
 ORGANISM
 Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Triticum.
 REFERENCE
 1 (bases 1 to 681)
 AUTHORS
 Wheat functional genomics - Glenlea developing seeds cDNA libraries
 TITLE
 Unpublished (2002)
 JOURNAL
 Contact: Dr. Sylvie Cloutier
 COMMENT
 Cereal Research Centre, Agriculture and Agri-food Canada
 195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
 Tel: (204) 983-2340
 Fax: (204) 983-4604
 Email: scloutier@em.agr.ca
 was cloned directionally, not all sequences generated with reverse
 primer were from the 5' end (same with forward primer and 3' end).
 Average insert size is >1.4 kb
 Plate: 017 row: G column: 12
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1. 681
 /organism="Triticum aestivum"
 /cultivar="Glenlea"
 /db_xref="taxon:4565"
 /clone="TAE15017G12R"
 /clone_lib="TAE15"
 /tissue_type="developing seeds"
 /dev_stage="15 days after anthesis"
 /lab_host="E. coli DH10B"
 note="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
 Site_1: NotI; Site_2: MluI; mRNA obtained from wheat seeds
 of cultivar Glenlea 15 days post-anthesis"
 BASE COUNT
 149 a 165 c 207 g 160 t

FEATURES source

Query Match 32.0%; Score 537.6; DB 14; Length 681;
 Best Local Similarity 86.9%; Pred. No. 3.8e-97;
 Matches 591; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 768 CAACCTGCTTGAAGCGCGGACGATGAGTGGTGCTGACTACATCTCTCTGTTGTTGA 827
 DB 1 CACCCCTGGCTCAAGCTGGGGATACAGTGGTGAGTCTCTCTTGGATTGA 60

QY 828 TAAAGTGATTCCTATATTCACCTCCGACAGGACAGTACCTCCGCTTCTTCTGCTGC 887
 DB 61 CTCCTGGATACCCACATCTCTGCTCCGACAGGACAGCTACCTACCTTCTGCTGC 120

QY 888 TGTTCAGATGCTTCTCCATCAGCGTGGTGATGAGTTCGCTGCTGCTGCTGCTGCTG 947
 DB 121 TGTTCAGATGCTTCTCTCCATCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180

QY 948 TGGCACCCTCAAGATTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1007
 DB 181 TGGCACCCTCAAGTTGGGACCCAGTCCGCTGCTGGCATAGGAGACTCCGAATGC 240

QY 1008 CACGCTCACTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1067
 DB 241 CACGCTCACTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300

QY 1068 TGTTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1127
 DB 301 TGTTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

QY 1128 AAAGCCTGGCTCTATCACACCCGACACCAAGTTTGAGGCTGTTGTATGTGCTTAGAA 1187

Db 361 AAAGCCTGGCTTCCATCAGCCACACCAAGTTTGGAGCTGTTGTATGTGCTCAAGAA 420
 QY 1188 GGAAGAGGGTGGCGGACACACTTCTTCTTCCCTGGTTACCGCCACAGTTCTACATGCG 1247
 Db 421 GGAGAGGGTGGCGGACACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480
 QY 1248 GACAACTGATGTGACAGGAGTGTGACTACGATTATGAATGACAAGGATGAGGAGCGAA 1307
 Db 481 GACTACTGATGTACAGGGGATGTGACAACATATGAATGACAAGGATGAGGAGCGAA 540
 QY 1308 GATGTGATGCTGCTGGTGACCGTATCAAAATGATTTAGCTCATCCAGCCTGTTCTGTTG 1367
 Db 541 GATGTGATGCTGCTGGTGACCGTATCAAGATGTTGTTGAGCTCATCCAGCCCTGGCTTG 600
 QY 1368 TGAGCAGGGTATGAGTTTCTGCTATCCGTGAGGTGGTAAAGCCTGGTCCCGTGTCTCAT 1427
 Db 601 TGAGCAGGGTATGAGTTTCTGCTATCCGTGAGGTGGTAAAGCCTGGTCCCGTGTCTCAT 660
 QY 1428 CAACAAATCATTCAGTAA 1447
 Db 661 CAATAATCATTCAGTAA 680

RESULT 5
 AV943931/c
 LOCUS
 DEFINITION
 AV943931 K. Sato unpublished cDNA library, strain H602 adult,
 heading stage top three leaves Hordeum vulgare subsp. spontaneum
 cDNA clone bahl9g17 3', mRNA sequence.
 AV943931
 VERSION
 AV943931.1 GI:18239728
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Hordeum vulgare subsp. spontaneum.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 REFERENCE
 1 (bases 1 to 736)
 AUTHORS
 Sato, K., Saisho, D. and Takeda, K.
 TITLE
 Barley EST sequencing project in NIG and Okayama Univ
 JOURNAL
 Unpublished (2002)
 COMMENT
 Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 Location/Qualifiers
 1. 736
 /organism="Hordeum vulgare subsp. spontaneum"
 /strain="H602"
 /db_xref="taxon:77009"
 /clone="bahl9g17"
 /clone_lib="K. Sato unpublished cDNA library, strain H602
 adult, heading stage top three leaves"
 /tissue_type="top three leaves"
 /dev_stage="adult, heading stage"
 BASE COUNT
 189 a 214 c 158 g 173 t
 ORIGIN

Query Match 31.1%; Score 523.4; DB 10; Length 736;
 Best Local Similarity 86.1%; Pred. No. 2.5e-94;
 Matches 617; Conservative 0; Mismatches 91; Indels 9; Gaps 3;

QY 903 CTCCTATCAGCGTCTGGTGTACAGTTCCACTGGCCGTATAGACCGTGGCACCGTCAAGAT 962
 DB 736 CTCCTATCAGCGTCTGGTGTACAGTTCCACTGGCCGTATAGACCGTGGCACCGTCAAGAT 677
 QY 963 TGTGTGACAGTGTATGCTGGAATCCGGACACCCGGAACCTGCACGCTCACTGGTGT 1022
 DB 676 TGGGGACCCAGTGCAGCTCTGTCGATCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 617

/clone="Sc01_11c12"
/clone_lib="Sc01_AAFc_ECORC_cold_stressed_winter_rye_seed
ings"
/tissue_type="leaf, crown"
/dev_stage="seedling three-leaf stage"
/notes="Vector: Bluescript SK-/XhoI-ECORI; Site.1: Eco RI;
Site.2: Xho I; Sampled three-leaf seedlings treated for
one week at 20C, 12 hrs light/day. Library made with
Stratagene UNIZAP XR Kit/Gigapack III Gold Kit. Lambda
library is amplified, then mass excised in SOLR cells."
BASE COUNT 194 a 227 c 178 g 35 others
ORIGIN
Query Match 31.0%; Score 521.2; DB 12; Length 761;
Best Local Similarity 80.1%; Pred. No. 6.8e-94;
Matches 613; Conservative 5; Mismatches 140; Indels 7; Gaps 2;
QY 794 ATGAGTGGTGCACATCTCTCGTTGGTGTGATAAGTGGATCTCTATATCCAGTCC 853
Db 759 ACGAGTGGTGGATCTCTCTCTGTGATTTGGCTCCGTCGATCCCATCCCTGTCC 700
QY 854 CGCAGAGCGACATGACCTCCCGTTCTTGTCTGCTGTTGAAGATGTTCTTCCATCACC 913
Db 699 CGCAGAGCGACACGACCTACCTCTTGTCTGCTGTTGAGGATGCTTCTCCATCACTG 640
QY 914 GTCGTGTACAGTTGGCAGCTGCGCTATAGAGCGTGCACCGCTCAAGATTGGTGACACAG 973
Db 639 GCGGTGTACAGTTGGCAGCTGCGCTATAGAGCGTGCACCGCTCAAGATTGGTGACACAG 580
QY 974 TCGATATCGTCGGATCCGGACACCGGAACTGCACCGTCACTGGTGTGAGATGTTCC 1033
Db 579 TCGAGCTCTGTCGGATCAGGAGACCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
QY 1034 AGAAGACCATGGATGATGCCATGGCCGAGAGCAATGTTGGGCTGCTGCTGCTGCTGCTGCT 1093-
Db 519 AGAAGACCATGGATGATGCCATGGCCGAGAGCAATGTTGGGCTGCTGCTGCTGCTGCTGCT 460
QY 1094 AGAAGATGACATGAAAGAGCGATGCTGCGCAAGAGCGTGGCTCTATCACACCGGACA 1153
Db 459 NGNNGGCGNTTGTAGNGGCGTGTGTTGGCAAGCGCGGTCCATCNGGCGNCA 400
QY 1154 CCAAGTTTGGAGCTGTTGTGTATGCTTAAAGAGGAGAGGTTGGCGGACACTCACCTT 1213
Db 399 CCAAGTTTGGAGCTGTTGTGTATGCTTAAAGAGGAGAGGTTGGCGGACACTCACCTT 340
QY 1214 TCTTCCCTGGTTACCGCCACAGTTCTACATGCGGACAACTGATGTCACAGGAGTGTGA 1273.
Db 339 TTTTCCCTGGTTACCGCCACAGTTCTACATGCGGCGCTGCTGCTGCTGCTGCTGCTGCT 280
QY 1274 CTACGATATGATGACAGGATGAGGAGGAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCT 1333
Db 279 CAACCATTTGAGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 220
QY 1334 AAATGATTTGTCAGCTCATCCAGCCTGTTGCTGTGAGCAGGAGTATGAGGTTTCTATCC 1393
Db 219 AGATGTTGTCAGCTCATCCAGCCTGTTGCTGTGAGCAGGAGTATGAGGTTTCTATCC 160
QY 1394 GTGAGGCTGTTAAGACCGTTGTCGCGGTCATCAACAATCATTCAGTAAACTGGAT 1453
Db 159 GTGAGGCTGTTAAGACCGTTGTCGCGGTCATCAACAATCATTCAGTAAACTGGAT 100
QY 1454 ATAACATATCCACCATGAGAAATTTCTTGTGTTTACTC-AAAGCGACATGCTCCGTAGTTG 1512
Db 99 GGAGGATATCCCGCTGAGATTTTCCCTNTTACTSTTTTGGGAAATGCTATGATGTTG 40
QY 1513 TTATTATGTTGAGTTTATGAGGTTGCTCATGTGCAATTTAGT 1557
Db 39 TTATTATGCG-----TTTAGGGAAGGCTCTTGTGAAATTTGTAGT 1

QY 1023 TGAGATGTTCCAGAAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
Db 616 TGAGATGTTCCAGAAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
QY 1083 CCGTGGTATGACAGGATGACATTTGAAAGAGGATGATGATGATGATGATGATGATGATGATGAT 1142
Db 556 CCGTGGTATGACAGGATGACATTTGAAAGAGGATGATGATGATGATGATGATGATGATGATGAT 497
QY 1143 CACACCGGACACCAAGTTTGGAGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1202
Db 496 CACGCCACACCAAGTTTGGAGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 437
QY 1203 ACATCTACCTTTCTTCCCTGGTACCGCCACACAGTTCTACATCGGACCAACTGATGTGAC 1262
Db 436 GCATCTCCCATTTTCCCTGGTACCGCCACACAGTTCTACATCGGACCAACTGATGTGAC 377
QY 1263 AGGAGTGTGACTACGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322
Db 376 GGGAGACGTGACAAACATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 317
QY 1323 TGACCGTATCAAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1382
Db 316 TGACCGTATCAAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 257
QY 1383 GTTGTCTATCGTGGAGTGTGTAAGACCGTTGTTGCGCGTGTCTATCAACAAAATCATTTGA 1442
Db 256 GTTGTCTATCGTGGAGTGTGTAAGACCGTTGTTGCGCGTGTCTATCAACAAAATCATTTGA 197
QY 1443 GTAACTGGATATACATATCCACCATGAGAAATTTTCTTGTGTTTACTC-AAAGCGACATG 1501
Db 196 GTAACTGGATATACATATCCACCATGAGAAATTTTCTTGTGTTTACTC-AAAGCGACATG 137
QY 1502 CTCCTAGTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1561
Db 136 CTCCTAGTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 83
QY 1562 CACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1618
Db 82 CACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 28
RESULT 6
BE705674/c
LOCUS Sc01_11c12 A Sc01_AAFc_ECORC_cold_stressed_winter_rye_seedlings 761 bp mRNA linear EST 12-SEP-2000
DEFINITION Secale cereale cDNA clone Sc01_11c12, mRNA sequence.
ACCESSION BE705674
VERSION BE705674.1 GI:10093939
KEYWORDS EST.
SOURCE rye.
ORGANISM Secale cereale
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Secale.
1 (bases 1 to 761)
AUTHORS Singh, J.A., Piche, C., Couroux, P., De Moors, A., Harris, L.J., Hattori
J.I., Ouellet, T., Robert, L.S., Spratt, D. and Tinker, N.A.
TITLE Expressed Sequence Tags from Cold-Stressed Winter Rye Seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Singh, J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
OC6, Canada
Tel: (613) 759-1562
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
FEATURES
1..761
Location/Qualifiers
source
/organism="Secale cereale"
/cultivar="Puma (winter rye)"
/db_xref="taxon:4550"

```

DEFINITION BJ308303 Y. Ogihara unpublished cDNA library, Wh_yd Triticum
aestivum cDNA clone whyd15g14 3', mRNA sequence.
ACCESSION BJ308303
VERSION BJ308303.1 GI:20116830
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 728)
Ogihara.Y. and Murai.K.
Expressed genes in Triticum aestivum
Unpublished (2002)
Contact: Tadasu Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..728
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone_lib="whyd15g14"
/tissue_type="spikelet at late flowering"
/dev_stage="Feekes' scale 6"
/Note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J. Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give phagescript phagemids in the TJ Close lab
at the University of California, Riverside (Akhunov, Chin
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."
BASE COUNT 194 a 212 c 153 g 169 t
ORIGIN
Query Match 30.9%; Score 520; DB 13; Length 728;
Best Local Similarity 86.3%; Pred. No. 1.2e-93;
Matches 635; Conservative 0; Mismatches 90; Indels 11; Gaps 5;
QY 881 TCCTCGCTGTTCAAGATGCTTCT-CCATCACCGGTCGTGTACAGTTGCCACTGG-CCG 938
Db 728 TCCTCGCTGTTGAGGATGCTTCTCCCATCACTGCTGTGTACAGTTGCCACTGGCCG 669
QY 939 TAGAGGGTGGCCACCGCTCAAGATTGGTGACACAGTCATATCGTGAATCCGGGACAC 998
Db 668 TATCAGCGTGGCACCCTGACAGTTGGGACCCAGTCACCTCGTCGCATCAGGAGAC 609
QY 999 CCGGAACCTGCAGGTCATCGTTGGTGTAGATGTTCCAGAAGACCATGGATGCCATGGC 1058
Db 608 TCGCAATGCCACCGGTCACTGGTGTGTAGATGTTCCAGAAGACCATGGATGCCATGGC 549
QY 1059 CGGAGACAATGTTGGGTGCTGCTCCGTGGTATCGACAGAGATGATGATGATGATGAT 1118
Db 548 TGGGACAATGTTGGGTGCTGCTCCGTGGTATCGAAGGAGACATTTGAGAGAGCAT 489
QY 1119 GGTGCTGCAAGAGCTGCTATACACCGCACACCAAGTTTGAGGCTGTTGTGTATGT 1178
Db 488 GGTGCTGCAAGAGCTGCTATACACCGCACACCAAGTTTGAGGCTGTTGTGTATGT 429
QY 1179 GCTTAAGAGAGAGGTTGCCACACTCACCTTTCCCTGGTTACCGCCACAGTT 1238
Db 428 CCTCAAGAGAGAGGTTGCCCGGCACTCCCAATTTTTCCTGGTTACCGTCGCAATT 369

```

```

QY 1239 CTACATGGCGGACAACTGATGTGACAGGAGTGTGACTACGATTATGAATGACAGGATGA 1298
Db 368 CTACATGGCGGACTACTGATGTGACGGGAATGTGACAAACATTATGAATGACAGGATGA 309
QY 1299 GGAGCGGAAGATGTGCATGCCCTGGTGACCGTATCAAAATGATTGTTCAAGTCTATCCAGCC 1358
Db 308 GGAGCGGAAGATGTGCATGCCGGGTGACCGTATCAAGATGTTGTGGAGCTCATCCAGCC 249
QY 1359 TGTTCCTTTGTGAGCAGGATATGAGCTTTGCTATCCGTGAGGTGGTAAGACCGTTGGTGC 1418
Db 248 CGTGGCTTTGTGAGCAGGGAATGAGCTTTGCCATCCGTGAAGTGGCAGACCGTCGGTGC 189
QY 1419 CGGTGTCATCAACAATAATCATTTAGTAACTGGATATAACATATCCACCATGAGAATTTT 1478
Db 188 CGGTGTCATCAATAATATCATTTAGTAACTGGATGGAGGATATCCACCGTGAGAATTTT 129
QY 1479 CCTTGTCTTACTC-AAAGCGACATGCTCGTAGTGTGTTATGTTTATGTTGAGTGTAGGGGT 1537
Db 128 CCTCATTTACTCTTTTGGGAATGCTATGATGTTGTTATTAAG-----CATTTAGGGA 75
QY 1538 TGCTCATGTGCAATTTAGTATGACACTTTTTTTTGTCAAGTGAATTTGTCATAATTTAT 1597
Db 74 GGCTCTTTGTGAATTTAGTATGACACTTTTTTCT--TCAGTGAATTTGCATCTTTGT 17
QY 1598 GACATTCACGACAAAG 1613
Db 16 AGTATTCACGACAAAG 1
RESULT 8
BG301275/c
LOCUS BG301275 900 bp mRNA linear EST 22-OCT-2001
DEFINITION HVSMEB0020G16f Hordeum vulgare seedling shoot EST library
HVCdNA0002 (Dehydration stress) Hordeum vulgare cDNA clone
HVSMEB0020G16f, mRNA sequence.
ACCESSION BG301275
VERSION BG301275.1 GI:13098802
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 900)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
,R.D., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex drought-stressed seedling shoot cDNA
library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hq bases = 221
Seq primer: AATTAACTCACTAAAGGG
High quality sequence start: 8
High quality sequence stop: 739.
FEATURES
Location/Qualifiers
1..900
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone_lib="HVSMEB0020G16f"
/HVCdNA0002 (Dehydration stress)
/tissue_type="Seedling shoot"
/lab_host="tjJC121"

```

/note="Vector: lambdaZAP; Site.1: EcoRI; Site.2: XhoI; Seeds were surface sterilized then germinated under axenic conditions in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Five-day old seedlings were incubated at 90° RH for 24 hr. Shoots were then harvested, total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, 600000 pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids. These steps were performed in the TJ Close laboratory at the University of California, Riverside (Choi, Close, Fenton). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see

<http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT
ORIGIN

225 a 258 c 212 g 205 t
Query Match 30.7%; Score 515.4; DB 12; Length 900;
Best Local Similarity 80.1%; Pred. No. 9.4e-93;
Matches 707; Conservative 0; Mismatches 161; Indels 15; Gaps 8;

QY 734 GCTCCGCTCAAGCGCTCGAGGCTCTCATGTGTCACCCCTGCTTGAAGCGGGGACG 793
DB 874 GGATGCACTCAAGCGCTCGAGGCTCTCAT-GCCACCCTGGCTCCAAGGGTGGGAATA 816
QY 794 ATGAGTGGGCGACTACATCTCTCGTTGGTGTGAATGGAATGCTATATCCAGTCC 853
DB 815 AAGAGTGGGTGGAGGC-TCTTTTCCCTGATGCTCCCTGACCACCCAAAT--CCTGTC 759
QY 854 CGCAGAGGCGAGCTGACCTCCCGTCTTCTGCTGCTGTTGAAGATGCTTCTCCATCCG 913
DB 758 CGCAAGGGCAGACAGACCTGCCCTTC-TGTTCTGCTGTAGGATGCTTTTCTCTCCCG 700
QY 914 GTCGTGGTAC-AGTTGCCACTGCCCTATAGACCTGGCCACCGTCAAGATGGTGACACA 972
DB 699 GTCGGGTCCCAAGTTGCCACTGACCTGACGCTGCGGACCGGTCAAGGTTGGGGACCA 640
QY 973 GTCGATATCTCGGAATCCGGGACACCGGAACCTGCACCGTCACTGGTGTGAGATGTTTC 1032
DB 639 GTCGACCTCTCGGCATCAGGAGACCCGCAATGCCACGGTCACTGCTGTAGATGTTTC 580
QY 1033 CAGAAGACCATGATGATGCCATGCCGCGGAGACAAATGTTGGGCTGCTGCTCCGTGGTATG 1092
DB 579 CAGAAGACCATGATGATGATGCTGCTGGGACAAATGTTGGGCTGCTGCTCCGTGATG 520
QY 1093 CAGAAGATCACATGGAAGGCGATGCTGCTGGGAAAGCCCTGGCTCTATCACACCGCAC 1152
DB 519 CAGAAGAACACATGAGATAGGCGATGGTGTCTCAAGCCGGGTTCCATCAGCCGACAC 460
QY 1153 ACCAAGTTTGAGGCTGTGTGTATGCTTGAAGGAAGAGGGTGGCGGCACACTCACCT 1212
DB 459 ACCAAGTTTGAGGCTGTGTATGCTCAAGAGGAGGAGGGTGGCGGCACCTCCCA 400
QY 1213 TTTTCCCTGGTTACCGCCACAGTCTTACATCGCGACAACTGATGTGACAGGAGTGTG 1272
DB 399 TTTTCCCTGGTTACCGTCCGAGTCTTACATCGCGACTACTGATGTCACGGGGAACGTG 340
QY 1273 ACTACGATTTATGATGACAGGATGAGGAGCGGCAAGATGTGATGCTGTCACCGTATC 1332
DB 339 ACAACATTATGATGACAAAGATGAGGAGGCGCAATATGTGATGCTCGGGTGACCGTATC 280

QY 1333 AAAATGATTGTTAGCTCATCAGCCTGCTGCTTGTGAGCAGGTGATGAGTGTTCCTATC 1392
DB 279 AAGATGGTGTGGAGCTCATCAGCCTGCTGCTTGTGAGCAGGAATAGGTTTGCCTATC 220
QY 1393 CGTGAAGGTGTAAGACCGTGTGCGGTGTCATCAACAAATCATTTAGCTAAACTGGA 1452
DB 219 CGTGAAGACGCGAAGACAGTCCGTGTCATCAACAAACATCATTCAGTAAAGTGA 160
QY 1453 TATAACATATCCACCATCAGAAATTTTCCTGTTTACTC-AAAGGCACATGCTCGTAGTT 1511
DB 159 TGGAGATATCCCGCTGAGAAATTTTCCTCATTTGCTGCTTTTGGGAATGCTATGATTT 100
QY 1512 GTTATTATGCTGAGTGTGAGGCTGCTCATGTGCAATTTGTAGTATGACACATTTTTTT 1571
DB 99 GCTATTATG-----GGTTTAGGAGAGCTCTTGTGAATTTGTAGTATGCTCTTC 46
QY 1572 TTGTCACATGAAATTTGCATAATTTATGACATTCACGACAAAGA 1614
DB 45 T--TCAAGTGACATTGCAAACTTTGTAGTATTCGCGGCAAAA 5

RESULT 9
BJ263237/c

LOCUS BJ263237 Y. Ogiwara unpublished cDNA library, wh_h Triticum
DEFINITION aestivum cDNA clone whh13j11 3', mRNA sequence.
ACCESSION BJ263237
VERSION BJ263237.1 GI:20084049
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum

REFERENCE
AUTHORS Ogiwara, Y. and Murai, K.
TITLE Expressed genes in Triticum aestivum
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..685
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whh13j11"
/clone_lib="Y. Ogiwara unpublished cDNA library, wh_h"
/tissue_type="spike at heading date"
/dev_stages="Feekes' scale 10.5"
Note: EORI; Site.2: XhoI; plants were grown under hydroponic conditions at UC Davis, salt stressed for 12 hours, and for 7 days, then dissected and frozen (Akhunov in J. Dvorak Lab). Total RNA was prepared from sheath tissue, equal quantities of RNA were pooled from the two samples, polyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab at the University of California, Riverside (Akhunov, Choi, Close, Fenton, Kianian, Otto, Simmons, Zhang). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 179 a 201 c 144 g 159 t
ORIGIN
Query Match 30.3%; Score 509.6; DB 13; Length 685;
Best Local Similarity 86.4%; Pred. No. 1.4e-91;
Matches 598; Conservative 0; Mismatches 86; Indels 8; Gaps 3;

RESULT 11

BI959687

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

BI959687 678 bp mRNA linear EST 22-OCT-2001
HVSMEn0020M20f Hordeum vulgare rachis EST library HVCDNA0015
(normal) Hordeum vulgare cDNA clone HVSMEn0020M20f, mRNA sequence.
BI959687
BI959687.1 GI:16310942
EST
Hordeum vulgare.
Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
1 (bases 1 to 678)
Wing, R., Close, R.J., Klein, H., A., Wise, R., Chin, A., Begum, D.,
Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons
J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex rachis cDNA library
Unpublished (2001)
Contact: Wing R
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Total hg bases = 490
Seq primer: AATTAACTCCTCACTAAAGG
High quality sequence stop: 556.
Location/Qualifiers
1..678
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEn0020M20f"
/clone_lib="Hordeum vulgare rachis EST library HVCDNA0015
(normal)"
/tissue_type="rachis"
/lab_host="TJCI21"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Plants were grown at Washington State University,
Pullman, WA in a greenhouse, the rachises were excised and
frozen in liquid nitrogen (Klein, H. lab). In the TJ Close
lab at the University of California, Riverside total RNA
was prepared, poly(A) was purified, one primary
unamplified cDNA library was made, and 1 million pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids
(Chin). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders
see Close TJ, Wing R, Klein, H. A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT 131 a 236 c 188 g 123 t
ORIGIN

Query Match 29.8%; Score 501.4; DB 13; Length 678;
Best Local Similarity 86.1%; Pred. No. 6.2e-90;
Matches 578; Conservative 0; Mismatches 91; Indels 2; Gaps 2;

QY 247 CCACACGTCAACATAGSCACCATCGGCATGTGCGACACGAGAAAGACCATCTCTCACCGCC 306

||||| 1 CCACACGTCAACCGCGCGACCATCGCCACAGTGCAGCACGCGCAAGCACCATCTCACGGCC 60
QY 307 GCGCTCACATGCTGCTCGCTCCCTCGGTGCGCAGCGCGCTTAAAGATAGACAGATC 366
||||| 61 GCGCTCACATGCTGCTCGCTCCCTCGGTGCGCAGCGCGCTTAAAGATAGACAGATC 120
QY 367 GACGCGCGCGCGAGGAGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 426
||||| 121 GACGCGCGCGCGAGGAGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 180
QY 427 GAGACGCGCGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 486
||||| 181 GAGACGCGCGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 240
QY 487 AATATGATCAACGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 546
||||| 241 AATATGATCAACGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 300
QY 547 GGGCCCATGCGCGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 606
||||| 301 GGGCCCATGCGCGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 360
QY 607 ATCGTTGCTTCTCTCAACAAAGAGACATGCTCGCGCGCGCGCGGTATCACCATCAAC 666
||||| 361 ATCGTTGCTTCTCTCAACAAAGAGACATGCTCGCGCGCGCGGTATCACCATCAAC 420
QY 667 GAGCTCGAGTCCGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 726
||||| 421 GATCTCGAGTCCGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 480
QY 727 GTCGCTGCGTCCGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 786
||||| 481 GTCGCTGCGTCCGCGCGCGCGCGCGGTATCACCATCAACACGCGCGCGTGGAGTAC 540
QY 787 GCGCAGCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 846
||||| 541 GGGGATAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
QY 847 CCAGTCCCGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 904
||||| 601 CCGTACCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 660
QY 905 CCATCACCGGT 915
||||| 661 CCATCACCGGT 671
RESULT 12
AV938814
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AV938814 644 bp mRNA linear EST 18-JAN-2002
AV938814 K. Sato unpublished cDNA library, strain H602 adult,
heading stage top three leaves Hordeum vulgare subsp. spontaneum
cDNA clone bahl917 5', mRNA sequence.
AV938814
AV938814.1 GI:18234611
EST
Hordeum vulgare subsp. spontaneum.
Hordeum vulgare subsp. spontaneum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
1 (bases 1 to 644)
Sato, K., Saisho, D. and Takeda, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadashi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae			
; Triticeae; Triticum.			
1 (bases 1 to 670)			
Ogihara,Y. and Murai,K.			
Expressed genes in Triticum aestivum			
Unpublished (2002)			
Contact: Tadasu Shin-i			
Center For Genetic Resource Information			
National Institute of Genetics			
1111 Yata, Mishima, Shizuoka 411-8540, Japan			
Tel: 81-559-81-6856			
Fax: 81-559-81-6855			
Email: tshini@genes.nig.ac.jp.			
location/Qualifiers			
1..670			
/organism="Triticum aestivum"			
/cultivar="Chinese Spring"			
/db_xref="taxon:4565"			
/clone="whf23el1"			
/clone_lib="Y. Ogihara unpublished cDNA library, wh_f"			
/tissue_type="spike at flowering date"			
/dev_stage="feekes' scale 10.5.1"			
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;			
Site.1: EcoRI; Site.2: XhoI; plants were grown under			
hydroponic conditions at UC Davis, salt stressed for 12			
hours, and for 7 days, then dissected and frozen (Akhunov			
in J Dvorak Lab). Total RNA was prepared from sheath			
tissue, equal quantities of RNA were pooled from the two			
samples, polyA was purified from the pooled RNA, a cDNA			
library was made, and the cDNA clones were in vivo			
excised to give pBluescript phagemids in the T7 Close lab			
at the University of California, Riverside (Akhunov, Chin			
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).			
Plasmid DNA preparations and DNA sequencing were			
performed in the OD Anderson lab (all other authors)."			
173 a 199 c 141 g 156 t 1 others			
BASE COUNT			
ORIGIN			
Query Match 29.6%; Score 498.4; DB 13; Length 670;			
Best Local Similarity 86.1%; Pred. No.2.5e-89;			
Matches 577; Conservative 0; Mismatches 86; Indels 7; Gaps -2;			
QY	904	TCCATCACCGGTCGTGGTACAGTTGCCACTGCCCTATAGACGTGGCCACCGTCAAGATT	963
Db	670	TCCATCACGTGGTGGTACAGTTGCCACTGGCCGTATCGACGTGGCCACCGTCAAGATT	611
QY	964	GGTGACACAGTCGATATCGTCGGAATCGGGACACCCGGAACCTGACAGGTGCTGGT	1023
Db	610	GGGGACCCAGTCGACTCGTCGGCATCAGGAGACTCGCAATGCCAGTGCACGTGGT	551
QY	1024	GAGATGTTCCAGAGACCATGGATGCCATGCCCGGAGACATGTTGGGCTGCTGCTC	1083
Db	550	GAGATGTTCCAGAGACCATGGATGCCATGCTGGGGAACAATGTTGGGCTTCTGCTC	491
QY	1084	CGTGGTATCCAGAGGATGACATTCAGAGAGCATGGTCTGGCAAGCCCTGGCTCATC	1143
Db	490	CGTGGTATCCAGAGGAGACATTCAGAGAGCATGGTCTGGCAAGCCCGTTCATC	431
QY	1144	ACACCGCACACCAAGTTTGAGGCTCTTGTGTATGTGCTTAAGAAGAGAGGGTGGCGA	1203
Db	430	ACGCCACACACCAAGTTTGAGGCTCTTGTGTATGTGCTCAAGAAGAGAGGGTGGCGG	371
QY	1204	CACCTACCTTCTTCCCTGGTTACCGCCACAGTTCTACATGCCGACACATGATGTGCA	1263
Db	370	CACCTCCCAATTTTTCCTGGTTACCGTCCGCAATTTCTACATGCCGACTACTGATGTCACG	311
QY	1264	GGGAGTGTGACTACGATTATGAATCACAAGATGAGAGGCGCAAGATGTGCATGCCTGGT	1323
Db	310	GGGAATGTGACAAACATTTAATGATCACAAGATGAGAGGCGCAAGATGTGCATGCCGGT	251
QY	1324	GACCGTATCAAAATGATTGTTCCAGCTCATCAGCCTGTTGCTTGTGACGAGGGTATGAGG	1383
Db	250	GACCGTATCAAGATGGTGTGGAGCTCATCCAGCCGCTGCTGTTGTGACGAGGAATGAGG	191

for the maize genome project. Sequences are present from libraries 947 and 949. Contigs were assembled using zmbAssembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."

BASE COUNT	156 a	154 c	106 g	142 t
ORIGIN				
Query Match	29.3%; Score 493.2; DB 13; Length 558;			
Best Local Similarity	93.8%; Pred. No. 2.8e-88;			
Matches 525; Conservative	0;	Mismatches 33;	Indels 2;	Gaps 1;
Qy 1048	GATGCCATGGCGGAGACAATGTTGGCTGCTCCGCTGATGACAGAGGATGACATT 1107			
Db				
Qy 558	GATGCCATGGCTGGAGACAATGCTGGACTGCTGCTCGTATGACAGAGGATGACATT 499			
Db				
Qy 1108	GAAAGAGGCATGGTGGTGGCAACCTGGCTCTATACACCGCACACCAAGTTGAGGCT 1167			
Db				
Qy 498	GAGAGAGGCATGGTGGCAAGCCGGCTCTATACACCGCACACCAAGTTGAGGCT 439			
Db				
Qy 1168	GTTGTGATGCTTAAGAAGAGAGGGTGGCGGACACTCACCTTCTCCCTGGTTAC 1227			
Db				
Qy 438	GTTGTGATGCTTAAGAAGAGAGGGTGGCGGACACTCACCTTCTCCCTGGTTAC 379			
Db				
Qy 1228	CGCCACACAGTTCTACATCGGACAACTGATGTGACAGGGAGTGTGACTACCAATTATGAAT 1287			
Db				
Qy 378	CGCCGCGCAGTTCTACATGAGCACAAACCGATGTGACAGGAATGTGACTGTGATTATCAAT 319			
Db				
Qy 1288	GACAAAGATGAGGAGGCGAAGATGTGCATGCTGTCGACCGTATCAAAATGATTGTTTAC 1347			
Db				
Qy 318	GACAAAGATGAGGAGGCGAAGATGTGCATGCTGTCGACCGTATCAAAATGATTGTTTAC 259			
Db				
Qy 1348	CTCATCCAGCCTGTTGCTTGAGCAGGGTATGAGGTTTGCTATCCGTGAGGGTGGTAAG 1407			
Db				
Qy 258	CTCATCCAGCCTGTTGCTTGAGCAGGGTATGAGGTTTGCTATCCGTGAGGGTGGTAAG 199			
Db				
Qy 1408	ACCGTTGGTGGCGGTGTCATCAACAAATCATTTGAGTAACTGGATATACATATCCACC 1467			
Db				
Qy 198	ACTGTTGGTGGTGAATAAACAACAAATCATTTGAGTAACTGGATATACATATCCACC 139			
Db				
Qy 1468	ATGAGAATTTCCCTGTTTACTCAAGCGACATGCTCCGTAGTTGTTATTATGTTGAG 1527			
Db				
Qy 138	ACGAGAATTTCCCTGTTTACTCAAGCGAATGCTCTGTAGTTGTTATTATGTTGAG 79			
Db				
Qy 1528	TTTTAGGGGTTGCTCATGTGCAATTTGATGACACTTTTTTTTTTTGTCAAAGTGAATTG 1587			
Db				
Qy 78	TTTTAGGGGTTGCTCATGTGCAATTTGATGACACTTTTTTTTTTTGTCAAAGTGAATTG 21			
Db				
Qy 1588	CATAATTTATGACATTCACG 1607			
Db				
Qy 20	CATAATTTATGACATTCACG 1			

Search completed: May 1, 2003, 10:18:51
Job time : 2907 secs

Db 787527 TCGCCCATCGACCAACAAGACACCGTCACCGGTTGGAGATGTTCCGCAAGCTGCTGA 787586
QY 1047 TGATGCCATGGCGGAGACAATGTTGGGCTGCTGCTCGGTATGCAAGAGGATGACAT 1106
Db 787587 CCAGGGCGAGGGCGGACAAAGTTGCTGCTGCGGGCGTCAAGCCGAGGACGT 787646
QY 1107 TGAAGAGCATGCTGCTGCGGAAGCTGGCTCTATCACACCGCACCAAGTTTGGAGC 1166
Db 787647 CGAGCGTGGCAGGTTGTCACCAAGCGCGGACCAACCGCCGACACCGAGTTCGAAG 787706
QY 1167 TGTGCTGATGCTTAAGAAGGAGAGGTTGGCGGACACTCACCTTCTTCCCTGGTTA 1226
Db 787707 CCAGGCTACATCTCTCAAGGAGAGGCGCGGACACACCGGTTCTTCAACAAC 787766
QY 1227 CGGCCACAGTTCTACATCGGCAACACTGATGTGACAGGAGTGTGACTACGATTATGAA 1286
Db 787767 CGGTCGAGTTCTTCCGACACCGAGTGTGACCGGTTGTTGAC----- 787814
QY 1287 TGACAAGATGAGGAGGCGAAGATGTGATGCTGACAGGAGTGTGACTACGATTATGAA 1346
Db 787815 ---ACTGCGGAGGCGACCGAGATGGTATGCCCGGTGACAACACCAACATCTCGGTGAA 787871
QY 1347 GCTCATCCAGCTGTTGCTTGTGACGAGGATGATGCTGCTGAGGTTGCTATCCGAGGTTGATA 1406
Db 787872 GTTGATCCAGCGGCTGCGCATGAGGAGGTTGCTGCTGCGGATCCGAGGTTGCGG 787931
QY 1407 GACGTTGCTGCGGTTGTCATCAACAATAATCATTTAGTA 1445
Db 787932 CACGTTGGCGCGCGGTTGTCACCAAGATCATCAAGTA 787970

RESULT 3

US-08-743-637B-185.
; Sequence 185, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: PICARD, Francois J.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 EAST WISCONSIN AVENUE
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,637B
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,840
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 185:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Haemophilus influenzae
; US-08-743-637B-185

Query Match 25.0%; Score 420.4; DB 2; Length 1185;
Best Local Similarity 60.8%; Pred. No. 5.1e-89;
Matches 747; Conservative 0; Mismatches 436; Indels 45; Gaps 2;

QY 219 GCGAGGGGCAAGTTGAGCGGCAACCAACACACGTCACATAGGCGCCCTCCGCTCGGTGG 338
Db 3 GTCTAAGAAAATTTGACGTACAAAACGACGTAAGCGTGGGTACAAATCGGCACGT 62
QY 279 CGACCAGGAAAGACCACTCTCACCGCGCGCTCACATGCTGCTCGCTCCGCTCGGTGG 338
Db 63 TGACCACGGTAAACAACCTTAAACAGCAGCAATCACAAACGTTATGGCAAAACATTACGG 122
QY 339 CAGCGGCTTAAGAAGTACAGAGATCGACGCGCCCGGAGGAGCGCGCGGTAT 398
Db 123 TGGTGCAGCGGTCGATTTCGACCAAAATTTGATAACGCGCCAGAGAAAAGCGGTGTAT 182
QY 399 CACCATCAACACCGCCACCGTCCAGTACGAGACCGAGACCCGCCACTACGCACACGTGA 458
Db 183 TACCATCAACACTTACAGCTTGAATACGATACACGCGCTCGCCACTATGCAACACGTGA 242
QY 459 CTGCCCCGCGCACGCGGCTATGTCAGAAATATGATCACGCGGCTGCGCAGATGACGG 518
Db 243 CTGTCGGGACACGCGGCTATGTTAAATATGATTACTGTCGCGCACAATGGATGG 302
QY 519 TGCATCTCTGCTGATCCGCTGCCGACGCGCCCATGCGCGACAGCAAAAGACACATCT 578
Db 303 TGTATTTTAGTAGTAGCAGCAACAGATGCTCTATCGCGCAACTCTGTAACACATCTT 362
QY 579 CTTGCGCAAGCAAGTCGGTGTCCCAAGATCGTTGCTTCTTCAACAACAAGACATGTT 638
Db 363 ATTAGTCCCAAGTAGTGTGCCATACATCATCTGTTCTTAAACAATGGACATGTT 422
QY 639 CGACGAGGAGGCTGCTCGAGCTCGAGCTCGAGGTCGCGGAGCTGCTCAGCAACTA 698
Db 423 AGATGACGAAGATTTAGTAATAGTAAATGGAAGTTCTGTAACCTTCTATCTCAATA 482
QY 699 CGAGTACGACGCGGACGAGTACCATCTGCTGCTCGGCTCGGCTCAAGGCGCTGAGGC 758
Db 483 TGACTTCCAGGTGACGATACACCAATCGTGTGCTTCCAGCATTAACAAGCGTTAAA --- 539
QY 759 TCTCATGGTCAACCCCTGCGCTTGAAGCGCGGACGATGATGCTGCTGCTGCTTCTTC 818
Db 540 -----CGCGGTAGCAGATGGGAAGAAAATCTCTGA 572
QY 819 GTTGGTTGATAAGTGGATTCTCTATATCCAGTCCCGCAGAGGACAGCTGACCTCCCGTT 878
Db 573 GTTAGCAAAACCACTTAGATACTTACATCCAGAACAGAAACGTCGATTGACCAACCGTT 632
QY 879 CTTGCTGCTGTTGAAGATGCTCTCCATCACCGCTGCTGCTGCTGCTGCTGCTGCTG 938
Db 633 CCTTCTTCCAATCGAAGATGTTCTCAATCTCAGGTGCTGCTGCTGCTGCTGCTGCTG 692
QY 939 TATAGAGCTGGCAGCTCAAGATTGGTGACACAGTTCGATATCTGCTGGAATCCGGGACAC 998
Db 693 TGTAGAACGAGGTATTATCCGTACAGGTGATCAAGTATGAGAAATCTGCTGCTATCAAGATAC 752
QY 999 CCGGAACCTCCAGGCTGCTGTTGATGATGTTTCAGAGACCATGAGATGATGATGATG 1058
Db 753 ACCGAAACTACTTAACGGGTGTTGAAATGTTCCGTAATTTACTTTGACGAAGGTCGTGC 812
QY 1059 CGGAGACATGTTGGCTGCTGCTCGGTGATGAGAGGATGACATTGAAAGAGGCAT 1118
Db 813 AGGTCAAAACATCGGTGCTATTTACGTGGTACCAACACGTTGAAGAAATCGAACGCTGTC 872


```

; CURRENT FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: staphylococcus aureus
US-09-218-197-1

```

Query Match	- 23.4%;	Score 393.2;	DB 4;	Length 1185;
Best Local Similarity	59.4%;	Pred. No. 1.le-82;		
Matches 730;	Conservative	0;	Mismatches 453;	Indels 45; Gaps
QY	219	GGCGGGGGCAAGTTCGAGCGCACCAAAACACACAGTCAACATAGGCACCATCGGCCATGT	278	
DB	3	GGCAAAAGAAATTCGATCGTTCTTAAGAACATGCCAATATCGGTACTATCGGTACAGT	62	
QY	279	CGACCAAGAAAGAACACTCTCACCGCGCGCTACCACTATGGTGTCTCGCTCCGTCGGTGG	338	
DB	63	TGACCATGTTAAACAACATTAACAGCAGCAATCTACTGTATTAGCAAAAATGGTGA	122	
QY	339	CAGCGCGCTAAGAGTAGCAGCAGATCGACGCCGCCCGCGAGGAGCGCGCCGGGTAT	398	
DB	123	CTCAGTTGCACAAATCATATGACATGATGACAACTCCAGAAGAAAAGAACGTGTTAT	182	
QY	399	CACCATCAACACCGGCACCGCTCGAGTAGCAGACCGGACCGCCACTACGCACACGTCGA	450	
DB	183	CACAATCAATCTTCTCACATTGAGTACCAAACTGACAAACGTCACTACGCTACAGTTGA	242	
QY	459	CTGCCCGCGCCACGCGGAGCTATGTCAAGAATATGATCACCGCGCTGCGAGATGGAGG	518	
DB	243	CTGCCCAGACACGCTGACTAGCTTTAAAACATGATCACTGGTGTCTCAAAATGGACGG	302	
QY	519	TGCCATCTCTGCTATCGGTCGCGACGGGCGCCATCGGCAGACCAAGACACATCCT	578	
DB	303	CGGTATCTTTAGTAGTATCTGCTGCTGAGCGTCCAAATGCCAACAACCTCGTGAACATCT	362	
QY	579	CCTCGCAAGCAAGTCGGTGTCCCAAGATCGTCTTCCTCAACAAGAGGACATGCT	638	
DB	363	TTTATCACCTNAGTTGGTGTACAGCATTAGTAGTATTCCTTAACAAGTTGACATGGT	422	
QY	639	CGAGCAGGAGAGTGTCTGAGCTCGTCGAGCTCGAGGTCCGCGAGTGTCTCAGCAACTA	698	
DB	423	TGACGATGAAGAATATTAGAAATTAGTAGAAATGAAAGTCTGTGACTTATTAAAGCAATA	482	
QY	699	CGAGTACGACGGCAGCAGGTACCAATCGTCGCTGGCTCCGCCCTCAAGGCGCTCGAGGC	758	
DB	483	TGACTTCCAGGTGACGATGTACCTGTAAATCGCTGGTTTCAGCATTTAAAGCTTTAGA	539	
QY	759	TCTCATGTGTCAACCTTGCCCTTGAAGCGGGCGACCATGAGTGGGTGCACTACATCTTCTC	818	
DB	540	-----AGGCGATGCTCAATACGAGAAAATCTTAGA	572	
QY	819	GTGTGTTGATAAGTGGATTCCTATATTTCCAGTCCCGCAGAGGACAGTACCTCCCGTT	878	
DB	573	ATTAATGGAAGCTGTAGACTTACATTTCCAACTCCAGAAGCTGATCTCGACAAACCAAT	632	
QY	879	CTTGCTCGCTGTTGAAGATGCTCTCCCATCACCGTCTGGTACAGTTGGCCATGGCCG	938	
DB	633	CATGATGCCAGTTGAGGACGTATTCCTCACTCACTGGTGGTACTGTGTGTACAGCGCG	692	
QY	939	TATAGAGCTGCGCCGCTCAAGATTGGTGACACAGTCCGATATCGTCGGAATCCGGGACAC	998	
DB	693	TGTTGAACGTGCTCAATCAAGTTGGTGAAGAAGTTGAATCATCGTTTACATGACAC	752	
QY	999	CGGAACTGCAGCTGCTGTTGAGATGTTCCAGAAGACCATGGATGATGCCATGGC	1058	
DB	753	ATCTAAAACAACCTGTACAGGTTGTTGAAATGTTCCGTAAATATTAGACTACGCTGAAGC	812	
QY	1059	CGGACACATGTTGGCTGCTGCTCCGTGGTATGCAGAAGAGATGACATTTGAAGAGGCAT	1118	
DB	813	TGCTGACACATTTGTTGATATTACAGTGGTGTGCTCGTGAAGACGCTACAACTGGTCA	872	

Qy	1119	GGTGTCTGGAAAGCCCTGGCTCTATACACCGCACACCAAGTTTGAGGCTGTTGTGTATGT	11718
Db	873	AGTATTAGCTGCTTCCCTGGTTCAATTACACCATATCTGAATTTAAAGCAGAAGTATACGT	932
Qy	1179	GCTTAAGCAAGGAAGAGGGTGGCGACACTCAGCTTCTTCCCTGCTTACCGCCACACAGTT	1238
Db	933	ATTATCAAAAGACGAAGGTGGAGTCCACACTCCATCTCTCAAACTATCGTCCACAAATT	992
Qy	1239	CTACATCGCGACAACTGATGTGACAGGAGTGTGACTAGCATTATTAATGACAGAGTGA	1298
Db	993	CTATTCCGTACTACTGACGTAACTGGTGTGT-----TCATTTACCAGA	1037
Qy	1299	GGAGGGCAAGATGTGATGCTCGTGGTGACCGTATCAAAATGATTGTTCAAGTCAATCCAGCC	1358
Db	1038	AGGTACTGAAATGGTAATGCTGGTGATAACGTTGAAATGACAGTAGAATTAATCGCTCC	1097
Qy	1359	TGTTGCTTGTGACACGGGTATGAGGTTGCTATCCGTGAGGGTGGTAAGACCGTTGGTGTC	1418
Db	1098	AATCGCGATTGAAGACGGTACTCGTTTTCATATCCGGAAGGTGGACGTACTGTAGGATC	1157
Qy	1419	CGGTGTCTCAACAAAAATCATTGAGTAA	1446
Db	1158	AGCGTGTGTACTGAAATCATTAATAAA	1185

RESULT 6

```

US-08-961-527-260
; Sequence 260, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 260:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2996 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-260

```

Query Match 21.9%; Score 368.8; DB 4; Length 2996;
Best Local Similarity 59.5%; pred. No. 7.8e-77;
Matches 706; Conservative 0; Mismatches 432; Indels 48; Gaps 3;

Db 80 CTTAAAGACTATGCGTCTATGATGCTGCTCCAGAGAAGACCGGATCATCACTATC 139
 Qy 406 AACACGCCACCGTCGAGTACGAGACGAGACCGCGCACTACGACACGCTGACGCCCC 465
 Db 140 AACACTGGCAGCTTGAGTACGAACATGAAAAGCTCACTACGCTACATCGAGCTCCA 199
 Qy 466 GGCACGCCGACTATGTCAGAATATATGATCACCGCGCTCGGAGATGGAGCGTCCATC 525
 Db 200 GGACACGGGACTACGTTAAACACATGATCACTGCTGCTCAATGGAGCGAGCTATC 259
 Qy 526 CTCGTCGATCCGTCGCGGAGCGGCCATGCGCGAGACCAAGAGACATCTCTCCGCC 585
 Db 260 CTTGTAGTACCTCAACTGAGCGACCAATGCCAATACTCGTGAACATCTCTCTTCA 319
 Qy 586 AAGCAAGTCGCTGTTCCCAAGATGTTCTCTCAACAAGAGGACATGTCGAGCAG 645
 Db 320 CGTCAGTGTGTTAAACACCTTATCGTCTTCAAGCAAAAGTGAATGTTGAGTGG 379
 Qy 646 GAGGAGCTGCTCGAGCTCGTGAAGTTCGAGTCCGCGAGCTGCTCAGCAACTACGAGTAC 705
 Db 380 GAAGAATTGCTTGAATGTTGAATGAAATCGTGACCTATTTGTCAGAATACGACTTC 439
 Qy 706 GACGGCAGCAGTACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
 Db 440 CCAGTGACGATCTTCCAGTTATCCAAAGTTTCAAGCTTAAAGCTTTGAAGGT----- 493
 Qy 766 GTCRAACCTGCTTGAAGCGGGGAGCAGTGAAGTGGTGCAGTACATCTTCTGTTGTT 825
 Db 494 -----GATCTAAATACGAAGACATCGTTATGGAATGATG 529
 Qy 826 GATAAGTGGATTCCTATATCCAGTCCCGCAGAGCAGTACGCTCCGCTTCTTCTC 885
 Db 530 AACACAGTTGATGATATATCCAGAACCAAGACGTGACACTGACAAACCATCTTCT 589
 Qy 886 GCTGTTGAAGATGCTTCTCCATCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
 Db 590 CCAGTGCAGGACGATTTCTCAATCACTGCGAGTGTGCTGCTGCTGCTGCTGCTGCT 649
 Qy 946 CGTGACCGCTCAGATGTTGTCACAGTGCATATCGTCGGAATC---CGGACACCCGG 1002
 Db 650 CGTGATCGTTAAGTCAAGCAGCAATCGAATCGTTGATCAAGAGAACTCAA 709
 Qy 1003 AACTGCAGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1062
 Db 710 AAAGCAGTGTGTTACTGCTGTTGAAATGTTCCGTAACAACTGACGAAGCTTGTCTG 769
 Qy 1063 GACATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1122
 Db 770 GATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
 Qy 1123 CTGGCAAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1182
 Db 830 ATCGTAAACAGGTTCAATCAACCCACACACTAAATCAAGGTGAAGTCTACATCCT 889
 Qy 1183 AAGAGGAAGAGGTTGGCCGACATCACTCTTCTTCCCTGTTTACCGCCACAGTTCTAC 1242
 Db 890 ACTAAGAAGAAGTGGACGCTCACTCTTCTTCAACAACTACCGCTCCACAAATCTAC 949
 Qy 1243 ATGGGCAACTGATGTCAGAGGAGTGTGACTACGATTAATGATGACAGGATGAGG 1302
 Db 950 TTCCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 994
 Qy 1303 GCGAGATGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362
 Db 995 ACTGAATGGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
 Qy 1363 GCTTGTAGCAGGATGATGAGTTGCTATCCGTCAGGCTGCTGAGACCGTTGGTGGCGGT 1422
 Db 1055 GCGCTAGAACAGGTACTACATTTCTTCTATCCGTCGAGGCTGAGCTACTGTTGTTCA 1114
 Qy 1423 GTCATCAACAAATCATTTGATGAATGATGATATCAATATCCACATGAGAAATTTCCCT 1482
 Db 1115 ATGGTTACAGAAATCGAAGCTTAATTCGATTTAGTTCCCGAGAACAAATTAATTA 1174

Qy 1483 GTTACTAAAGCAGCATGCTCCGTAGTTGTTATTATTATGTTGAGT 1528
 Db 1175 AGACACTAAAGAATCTTCTGTTGGCAAGGTTCTTTTATTAGATATT 1220

RESULT 7
 US-09-221-017B-642
 ; Sequence 642, Application US/09221017B
 ; Patent No. 6444799
 ; GENERAL INFORMATION:
 ; APPLICANT: Ross, Bruce C.
 ; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
 ; NUMBER OF SEQUENCES: 1120
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 PAGE MILL ROAD
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FastSeq for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/221,017B
 ; FILING DATE: 23-DEC-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP1182
 ; FILING DATE: 31-DEC-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP1546
 ; FILING DATE: 30-JAN-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP2911
 ; FILING DATE: 09-APR-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/AU98/01023
 ; FILING DATE: 10-DEC-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Monroy, Gladys H
 ; REGISTRATION NUMBER: 32,430
 ; REFERENCE/DOCKET NUMBER: 27340-20021.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-813-5600
 ; TELEFAX: 650-494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 642:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2592 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: UNKNOWN
 ; ORIGINAL SOURCE:
 ; ORGANISM: PORPHYROMONAS GINGIVALIS
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1...2592
 ; US-09-221-017B-642

Query Match 14.5%; Score 243.8; DB 4; Length 2592;
 Best Local Similarity 60.9%; Pred. No. 1e-47;
 Matches 432; Conservative 0; Mismatches 247; Indels 30; Gaps 1;

Qy 219 GCGAGGGGCAAGTTCTGAGCGCACCAACACACGTCACATAGGACCATCGGCATGT 278
 Db 1914 GCGAAGAGCATTTTAACAGATCGAACCACCCACGTTACGTTGGTACGATCGGACG 1973

Qy 279 CGACCACGAAAGACCACTCTCACCGCGCGCTCACCATGGTGTCTCCCTCCGTCGGTGG 338
Db 1974 GGACCAGGTAAGACTACCTTGACGGCTGCAATCACAACTGTGTGCAAGAGGGGTCT 2033
Qy 339 CAGCGCGCTAAGAGTACGACGAGATCGACGGCGCCCGAGAGAGCGCGCGCGGTAT 398
Db 2034 TTCAGAACTCCGTTTCATTGTGATCAATGATAAGCTCCCGAAGAAAGAAAGCTGGTAT 2093
Qy 399 CACCATCAACACCCCGCTGAGTACGAGACCGGAGACCGGCTACACACACGCTCGA 458
Db 2094 CACGATTAACTTCACACGTTGAATACCAACAGCTAATCGTACCTACGCTCACGTFAGA 2153
Qy 459 CTGCCCCGGCCACCGCACTATGTCAAGAATATGATCACCGGGCTCGCGAGATGGAAGG 518
Db 2154 CTGTCCGGGTCAACCGGCTATGTGAAGACATGTTACCGGTGCTGCTCAGATGAGCG 2213
Qy 519 TGCATCTCTGCTATCCGGTCCGACGGCGCCATCGCGACGACCAAGAGACATCTCT 578
Db 2214 TGCTAATCTGTTGACGAGTACAGACGGCGCTATGCCCTCAGACTCGCGAGCACATCTCT 2273
Qy 579 CTGCGCAAGCACTCGGTGTTCCCAAGATCGTGTCTCTCAACAAGAGGACATGCT 638
Db 2274 TTTGGCTCGTACGCTAAAGCTTCTCTGCTGTTGTTTTCATGACAAATGTGACATGGT 2333
Qy 639 CGACGACGAGAGCTGCTCGAGCTGCTGAGCTCGAGGTCCGGGAGCTGCTCAGCAACTA 698
Db 2334 AGACGATGAGAGATGCTCGAGCTGTTGTAATGGACATCGCGCAACTCTCTTCTTCTA 2393
Qy 699 CGACTACGCGGCGACGACGTACCAATCGCTGCTGGCTCGCGCTCAAGCGGCTCGAGGC 758
Db 2394 CGATTTCGATGGTACAAATACCCCTATCATCCGTTGCTGCTGGCGCTTT----- 2447
Qy 759 TCTCATGGTCAACCTCGCTTGAAGCGCGGACGATGAGTGGTGCAGCTACATCTCTCTC 818
Db 2448 -----GAATGGAGACCTCAGTGGGAGACAGGTGATGA 2483
Qy 819 GTTGGTGAATAGTGAATCTTATATCCAGTCCCGCAGAGGACGACTGACCTCCCGTT 878
Db 2484 GCTTATGGAAGCTGTTGACAACTGGGTTCCTCCCTGCCAGAGCGGATATCGACAAACGGT 2543
Qy 879 CTTCCTCGCTGTTCAAGATGCTTCTCCATCACGGGCTGCTGATGATGTT 927
Db 2544 CTTGATCCGGTTGAAGACGTGTTCTCTATCACGGGTGCTGGTACGGTT 2592

RESULT 8

US-09-140-466-1
; Sequence 1, Application US/09140466
; Patent No. 6268160
; GENERAL INFORMATION:
; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; FILE REFERENCE: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: US 60/056,246
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1230)
US-09-140-466-1

Query Match 11.2%; Score 188.8; DB 4; Length 1230;
Best Local Similarity -48.5%; Pred. No. 5,3e-35;
Matches 585; Conservative 0; Mismatches 612; Indels 9; Gaps 2;
Qy 239 GCACCAAAACACACAGTCAACATAGGCAACCATGGCCATGTGCAACACGGAAGACACATTC 298
Db 23 GAAATAAACAACATATAAATTTAGGTACTATAGGGCATGTAGATCATGTGAAAAAATACAT 82
Qy 299 TCACCCCGCGGCTCACCATGGTGTCTCCCTCGTGGTGGCAGCGGCTTAGAAGTAGC 358
Db 83 TAACAACAGCTATATCTTATTTAAATTTACAAGGATTATCAAAAAAATAATATTATT 142
Qy 359 ACAGATCGACCGCGCCCGGAGGAGCGCGGCTATCACCATCAACACCGCCACCG 418
Db 143 CAGATATTGATTCAGCTCCAGAGAAAAAATAAGAGGTATTACAATTAATACACACATA 202
Qy 419 TCGAGTACGAGACCGGACCGCCACTAGCGACACGTGCTGCTGCTGCTGCTGCTGCTGCT 478
Db 203 TTGAATATGAACTTTAAACAACATTTGCTCATATAGATTGCTCCAGGACATTCGGATT 262
Qy 479 ATGTCAAGAAATATGATCACCGGCTGCGCAGATGGCAGGTGCCATCTCGTCTCGT 538
Db 263 ATATTAAAAATATGATTATAGGAGCCACACAATGGATATACGAATTTTGTATATATCTA 322
Qy 539 GTGCGACGGGCGCCATGCGCGACAGCAACAAAGACACATCTCTCGCAAGCAAGTCGGTG 598
Db 323 TAATAGATGTTATGCTCAAACTTATGAACATTTATTATTAAATAAACAATAAGSTA 382
Qy 599 TTCCCAAGATGTTGCTTCTTCAACAAGAAGACATGGTGCAGCAGGAGAGTCTGCTCG 658
Db 383 TAAAAAATAAATATTTTAAATAAAGAAGATTATGTGATGATGTTGAATTAATAG 442
Qy 659 AGTCTGCTGAGTCCGAGGTCCGCGAGCTGCTCAGCAACTACGAGTACGAGCGGACGAG 718
Db 443 ATTTTATAAATTAGAAGTAATGAATTAATTAATTAATTAATTTGATTTAAATTTATA 502
Qy 719 TACCAATCGTGTGCTCGCGCTCAAGCGCTCGAGGCTCTCATGTGTCACCCCTGCT 778
Db 503 TACATATATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
Qy 779 TGAAGCGCGGCGACGATGAGTGGTGGTACATCTCTCTGCTGCTGCTGCTGCTGCTGCT 838
Db 563 ATGAATTAATAAATCTAATATTTGGATACAAAAATTAATAATTTAATTTCAAAATAAT 622
Qy 839 CTTATATCCAGTCCCGCAGAGCGACACTGACCTCCGCTTCTGCTGCTGCTGCTGCTGCT 898
Db 623 ATAAATATTAATACCTACTAGAAAAATTAATGATTACTTTTAAATGCAATAGAGATG 682
Qy 899 TCTTCTCCATCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958
Db 683 TATTTCTATAACAGGTAGAGGTACAGTAGTAACAGGTAAAGATTGAACAAGGATGTATA 742
Qy 959 AGATTGGTGACACAGTCGATATCGTGGAAATCGGGACACCC-----GGAAGTGCACGG 1012
Db 743 ATTTAAATGATGAAATGAAATTTTAAATTTGAAAAATTCATCTCTCTAATTTTAAACAAG 802
Qy 1013 TCAGTGGTGTGAGATGTTCCAGAGAGACCATGGATGATGCCATGGCCGAGCAATGTTG 1072
Db 803 TTATAGGATTAAGATGTTTAAACAACATTAACAACAGCAACATCCGGAGATATAGTAG 862
Qy 1073 GGTGCTGCTCGGTGATGTCAGAGAGGATGACATTTGAAAGAGGATGCTGCTGCTGCTGCT 1132
Db 863 GTATTTTAAAGAAATATTTCAAAAAAAGATATAAAAAAGAGGTATGATTTTGAACAAC 922
Qy 1133 CTGGCTCTATCACCGGACACCAAGTTGAGGCTGTTGAGTGTGATGCTGCTGCTGCTGCTGCT 1192
Db 923 CTAATAAATTAAGATATATAAGTCTTTTATAGCTCAAAACATATATTTTAACTAAAGAG 982
Qy 1193 AGGTGGCGGACACTCACCTTCTTCTCGCTGCTTACCGCCACAGTTCTACATCGCGGACAA 1252
Db 983 AAGTGGTGTCTATTAACCTTTTATATTTGATATTAACCTCAATTTTTTATTCGTACAG 1042
Qy 1253 CTGATGTGACAGGGAGTGTGACTAGCATTTATGAATGACAAGGATGAGGAGCGAAGATGT 1312

Db 1043 TAGATGTTACTG---GAGAAATTAATAATATATATTTAAATGAAATGTACAAAAGTAG 1099
QY 1313 GCATCCCTGGTGACCGTATCAAAATGATTTTCAGCTCATCCAGCCCTGTCCTTGTGAGC 1372
Db 1100 CTATACCTGGAGATAAAATACATTTACATATTTGAATTAATAACATATATAGTTGTGACAT 1159
QY 1373 AGGTTATCAGGTTTCTATCCGTGAGGTGTTAAGACCGTTGGTGGCGGTGTCAATCAACA 1432
Db 1160 TAAATATGAATTTCTATTAGAGAGGAGGAAAAACAATAAGGACGAGGTATTATATACAG 1219
QY 1433 AAATCA 1438
Db 1220 AAATAA 1225

RESULT 9

US-08-920-812-10/c
; Sequence 10, Application US/08920812
; Patent No. 5763188
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,812
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 3719 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecalis
STRAIN: Clinical Isolate S2-3
US-08-920-812-10

Query Match 10.8%; Score 182.2; DB 1; Length 3719;
Best Local Similarity 58.2%; Pred. No. 2.8e-33;
Matches 386; Conservative 0; Mismatches 258; Indels 19; Gaps 3;

QY 787 GCGGACGAGTGGTGGCGACTACATCTTCTCGTTGGTTGATAAGTGATTCCTATATT 846
Db 3708 GCGGACGAGTCTTATGAGAAAAAATCTAGAAATTAATGGCTGCAGTTGACGAATATATC 3649

QY 847 CCAGTCCCGCAGGCGAGCATGTACCTCCGTTCTTCTCGCTGTTGAAGATGTTCTCTCC 906
Db 3648 CCAACTCCAGACGTGATCTGACAAACCATTCATGATGCCAGTGAAGACGTTATCTCA 3589
QY 907 ATCACCGGTGGTGTACAGTTGCCACTGGCCGCTATAGAGCGTGGACCGCTCAAGATTTGT 966
Db 3588 ATCACTGGAGCTGGTACTGTTGTACAGGCGGTGTTGAACGTTGTTGAAGTTCCGCGTTGT 3529
QY 967 GACACATCGATATCGTGGAAATCCGGGAC---ACCCGGAACTGCGACGCTCACTGGTGT 1023
Db 3528 GACGAAATTTGAATCGTTGGTATTAAGAGCAACATCTAAACAACTGTTACAGGTGT 3469
QY 1024 GAGATGTTCCAGAAACCATGGATGATGCCATGGCCGGAGACAATGTTGGCTGTGCTC 1083
Db 3468 GAAATGTTCCGTAATTAATTAGACTAGCTGAAGCAGGCGACAACATCGTGTCTTTATTA 3409
QY 1084 CGTGGTATGCAGAAAGATGACATTTGAAAGAGGCGATGCTGGCAAAGCCTGGCTCTATC 1143
Db 3408 CGTGGTGTTCACGCTGAAGATATCGAACGTTGGACAAGTATTAGCTAAACGAGTACAATC 3349
QY 1144 ACACCGCACACCAAGTTTGAGCGTGTGTATGCTCTTAAGAAGGAAGAGGTTGCCGA 1203
Db 3348 ACTCCACACACAAAATTCAAAGCTGAAGTATACGTATTATCAAAAGAGAGGCGGACGT 3289
QY 1204 CACTCACTCTTCTCCCTGTTTACCCTGCTTACATCGCGACAACTGATGTGACA 1263
Db 3288 CACACTCA-TTCTTCACTAATCTGCTCTCAATTTCTACTTCGTTACACAGACGTTACT 3230
QY 1264 GCGAGTGTACTAGCATTTAATGACAAAGGATGAGAGGCGGAGAGTGTGCATGCTGTGT 1323
Db 3229 GGTGTGT-----AGAAATGCCCAAGGACTGAAATGTTAATGCTGTGT 3185
QY 1324 GACCGTATCAAAATGATTGTTACGCTCAATCCAGCCTGTTGCTGTGAGCAGGATGAGG 1383
Db 3184 GATAAGCTTGCTATGAGCAGTTGAATTAATCAACCAATCGCTATCGAAGACGAACTCTGT 3125
QY 1384 TTTGCTATCCGTGAGGTTGTTAAGACCGTTGGTGGCGGTGTCTCATCAACAAAATCATTTAG 1443
Db 3124 TTCTCTATTCTGTAAGCGGACGCTACTGTAGTTTCAGGCGTGTCTTACTGAAATCGTTAAA 3065
QY 1444 TAA 1446
Db 3064 TAA 3062

RESULT 10

US-08-920-827-10/c
; Sequence 10, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435

Db 3064 TAA 3062

RESULT 11
US-08-921-177-10/c
: Sequence 10, Application US/08921177
: Patent No. 5798211
: GENERAL INFORMATION:
: APPLICANT: Ohno, Tsuneya
: APPLICANT: Matsuhisa, Akio
: APPLICANT: Uehara, Hirotosugu
: APPLICANT: Eda, Soji
: TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402

```

/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/921,177
/ FILING DATE: 29-AUG-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION INFORMATION:
/ APPLICATION NUMBER: US 08/362,577
/ FILING DATE: 27-MAR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rin-Laures, Li-Hsien
/ REGISTRATION NUMBER: 33,547
/ REFERENCE/DOCKET NUMBER: 19036/32420
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6300
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3719 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecalis
/ STRAIN: Clinical Isolate S2-3
/
/ US-08-921-177-10
/
Query Match 10.8%; Score 182.2; DB 1; Length 3719
Best Local Similarity 58.2%; Pred. No. 2.8e-33;
Matches 386; Conservative 0; Mismatches 258; Indels 19
/
QY 787 GCGGCGAGTGGTGGCTGCACCTACATCTCTCGTTGGTTGATAAAGTGGAATCCCTCAT
/
Db 3708 GCGGCGAGTCTTATGAAGAAAATCTTAGAATTAATGGCTGCAGTTCACGAATAT
/
QY 847 CCAGTCCCGCAGGAGGACACTGACCTCCCGTTCTTGCTCGCTGTGTGAAGATGCTCTT
/
Db 3648 CCAACTCCAGAAGGTGATACACAAACCATTTCATGATGCCAGTCCAGACAGTAT
/
QY 907 ATCACGGTCTGGTACAGTTGCCACTTGGCCGCTATAGAGCGTGGGACCGTCAAGAT
/
Db 3588 ATCATGGACGTGGTACTGTTGCTACAGGCCGTGTTGAACGTTGGTGAAGTTCGCTCG
/
QY 967 GACACAGTCGATATCGTCGGAATCCGGGAC---ACCCGGAACTGCACCGTCACTGG
/
Db 3528 GACGAAGTTGAATCGTTGGTATTAAAGACGAACACATCTAAAACAACTGTTACAGG

```

[illegible]

STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/920.828
 FILING DATE: 29-AUG-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/362,577
 FILING DATE: 27-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Rin-Laures, Li-Hsien
 REGISTRATION NUMBER: 33,547
 REFERENCE/DOCKET NUMBER: 19036/32420
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3719 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecalis
 STRAIN: Clinical Isolate S2-3
 US-08-920-828-10

Query Match 10.8%; Score 182.2; DB 2; Length 3719;
 Best Local Similarity 58.2%; Pred. No. 2.8e-33;
 Matches 386; Conservative 0; Mismatches 258; Indels 19; Gaps 3;

QY 787 GCGACGATGAGTGGTGGCTACATCTCTCGTTGGTGTATGAAGTGGATTCCTATATT 846
 Db 3708 GCGACGAGTCTTATGAAGAAATCTTAGAATTAATGCTGCAGTTGACGAATATATC 3649
 QY 847 CCACTCCGCGAGCGAGACTGACCTCCGTTCTTGTCTGCTGTGAAGATGCTTCTCC 906
 Db 3648 CCAACTCCAGAACGTGATCTGACAAACCATTCATGATGCCAGTGAAGACGTATTCTCA 3589
 QY 907 ATCACCCTGCTGCTACAGTCCACCTGCGCGGTATAGAGCGTGGCACCGTCAAGATTGGT 966
 Db 3588 ATCACTGGACGTGCTACTGTGTGTACAGCCGCTGTTGAACGTGGTGAAGTTCGCGTTGGT 3529
 QY 967 GACACGTGATATCGTGGAAATCCGGGAC---ACCGGAACTGCAGGTCCTGCTGTT 1023
 Db 3528 GACCAAGTTGAATCGTGTGTTATTAAGACGAACATCTTAACACACTGTACAGGTGT 3469
 QY 1024 GAGATGTTCCAGAGACCATGGATGATGCGCGGAGACAAATGTGGGCTGCTGCTC 1083
 Db 3468 GAAATGTTCCGTAATATTATAGACTAGCTGAAGCGGCGACAAATCGTGTGCTTTATTA 3409
 QY 1084 CGTGGTATGCAAGGATGACATTGAAGAGCATGCTGCTGCAAGCGCTGCTCTATC 1143
 Db 3408 CGTGGTGTCCAGTGAAGATATCGAACGTGGACAGTATTAGCTAAACAGCTACAATC 3349
 QY 1144 ACACCGACACCAAGTTTGAAGCTGTGTGTATGCTTAAAGAAAGAGGAGGTCGCCCA 1203
 Db 3348 ACTCCACACACAAAATTCAAAGCTGAAGTATAGTATTATCAAAAGAGGCGGACGT 3289
 QY 1204 CACTCACCTTCTTCTGTTACCGCCCAACAGTCTTACATCGGACAACTGATGTGACA 1263
 Db 3288 CACACTCA-TTCTTCACTAACTATCGTCTCAATTCTACTTCCGTACAACAGACGTACT 3230

QY 1264 GGGAGTGTGACTACGATTATGAATGACAAGGATGAGGAGCGAAGATGTCATGCTCCTGGT 1323
 Db 3229 GGTGTGT-----AGAAATGCGAGAAGTACTGAATGGTAAATGCTCCTGGT 3185
 QY 1324 GACCGTATCAAAATGATTGTTCACTCATCCAGCTGTTGCTTGTGAGCAGGGGTATGAGG 1383
 Db 3184 GATAACGTTGCTATGGAGCTTGAATTAATTCACCCAATCGCTATCGAAGACGGAACCTGT 3125
 QY 1384 TTTGCTATCCGTGAGGTGGTGAAGACGCTTGGTCCCGGTGTCATCAACAAATCATTTGAG 1443
 Db 3124 TTCTCTATCGTGAAGCGGACGTACTGTAGTTTCAGCGCTTGTACTGAATCGTTAA 3065
 QY 1444 TAA 1446
 Db 3064 TAA 3062

RESULT 14

US-08-936-165A-87
 ; Sequence 87, Application US/08936165A
 ; Patent No. 6348582
 ; GENERAL INFORMATION:
 ; APPLICANT: Black, Michael
 ; APPLICANT: Burnham, Martin
 ; APPLICANT: Hodgson, John
 ; APPLICANT: Knowles, David
 ; APPLICANT: Lonetto, Michael
 ; APPLICANT: Nicholas, Richard
 ; APPLICANT: Pratt, Julie
 ; APPLICANT: Reichard, Richard
 ; APPLICANT: Rosenberg, Martin
 ; APPLICANT: Ward, Judith
 ; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
 ; TITLE OF INVENTION: Polypeptides and Their Uses
 ; NUMBER OF SEQUENCES: 534
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406-0939
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/936,165A
 ; FILING DATE: 24-SEP-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/027,032
 ; FILING DATE: 24-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gimmi, Edward R
 ; REGISTRATION NUMBER: 38,891
 ; REFERENCE/DOCKET NUMBER: P50549
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-270-4478
 ; TELEFAX: 610-270-5090
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 87:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 656 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Genomic DNA
 ; US-08-936-165A-87

Query Match 10.7%; Score 180.2; DB 4; Length 656;
 Best Local Similarity 62.4%; Pred. No. 4.3e-33;

Matches 295; Conservative 0; Mismatches 177; Indels 1; Gaps 1;
QY 286 GGAAGACCACTCTCACCGCGGCTCACCATGGTGTCCCTCCGTCGGTGGCAGCGG 345
Db 3 GGTAAACAACATTAACAGCAGCAATCGCTACTGTATTAGCAAAAATGGTACTCAGTT 62
QY 346 CCTAAGAGTACGACGAGATCGAGCGCGCCCGGAGAGCGCGCGGTATCACCATC 405
Db 63 GCACAATCATATGACATGATGTGACACGCTCCAGAGAAAAGAACGTTGATATCAATC 122
QY 406 AACACCGCCACCGTCGAGTACGACGAGACCGCCGCTACGACACGTCGACTGCCCC 465
Db 123 AATCTCTCANTTAGTACCACTGACAAAGTCACTANGCTCAGCTCGNCTGCCCA 182
QY 466 GG-CCACGCGGACTATGTCAGAAATATGATCACCAGCGGTCGCGAGATGACGGTGCAT 524
Db 183 GNAATCCGCTGACTACGTTAAACATGATCAGTGGTGTCTCAAAATGGACGCGGTAT 242
QY 525 CTTGCTGTATCCGGTCCGAGCGGCCCATGCGCGACGACCAAGAGCACATCCTCTCGC 584
Db 243 CTTAGTAGTATCTGCTGACGCTGCAATGCCAAGCTCGTGAACACATCTTTTATC 302
QY 585 CAAGCAAGTCGGTGTCCCAAGATCGTTGCTCTCTCAACAAGAGACATGTCGACGA 644
Db 303 ACGTAACGTTGGTACAGCATTAGTAGTATTCTTAAACAAGTTGACATGGTTGACGA 362
QY 645 CGAGGAGTGTCTGAGTCGTCGAGCTCGAGGTCGCGGAGCTGCTCAGCAACTACGAGTA 704
Db 363 TGAAGAATTATTAGAAATTAGTAGAAATGGAAGTTGCTGACTTATTAAGCGAATATGACTT 422
QY 705 CGACGCGCAGCAGTACCAATCGTCGCTGCTCCGCTCCGCTCAAGCGCTCGAGG 757
Db 423 CCCAGGTGACGATGTACCTGTAATCGCTGGTTCAGCATTAAGCTTTAGAG 475

RESULT 15
US-09-877-70/c
; Sequence 70, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556.877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 70
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-556-877-70
Query Match 6.7%; Score 112.6; DB 4; Length 323;
Best Local Similarity 60.7%; Pred. No. 1.9e-17;
Matches 184; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 879 CTTGCTCGCTGTGAAGATGCTCTCCCATCACCAGTGGTGTACAGTTGCCACTGGCGG 938
Db 323 CTTAATGCCCATTTAGAGCGTGTCTCTATCTCCGACGAGGAAGTGTAGTAACCTGGACG 264
QY 939 TATAGACGTGGCACCCTCAAGATTTGGTGACACAGCTGATATCGTCGGGAATCCGGGACAC 998
Db 263 TATTGACGTGGAATTTGTTAAAGTTTCCCGATAAAGTTCAAGTTGGTGGTCTTTAGAGATAC 204
QY 999 CCGGAAGTGCACGCTGCTGGTGTTCAGATGTTCCAGAGACCATGATGATGCCATGCC 1058
Db 203 TAAAGAAACGATTGTTACTGGGGTTGAAATGTTTCAAGAAAGAACTCCCAAGAGTCTGTCG 144

QY 1059 CGGAGACAATGTTGGCTGCTGCTCCGTTGGTATGCAAGAAGGATGACATTTGAAAGAGGCAAT 1118
Db 143 AGGAGAGAAATGTTGGATTGCTCCTCAGAGGTATTGGTAAGAAGCATGTGGAAAGAGGAAT 84
QY 1119 GGTGCTGGCAAGCCCTGCTCTATCACACCGGACACCAAGTTTGAGGCTGTTGTGTATGT 1178
Db 83 GGTGCTGTTGCTTGGCCAAACAGTGTAAACCTCATACAGGTTTAAAGTGTGCTGTTAGCT 24
QY 1179 GCT 1181
Db 23 TCT 21

Search completed: May 1, 2003, 11:17:53
Job time : 3545 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 15:58:02 ; Search time 273 Seconds
(without alignments)
7275.843 Million cell updates/sec

Title: US-09-810-764A-6

Perfect score: 1681

Sequence: 1 attcccaataatccccacc.....gttaaaaaaaaaaaaaa 1681

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.4	31.4	1194	10	US-09-815-242-7897
2	521.6	31.0	1194	10	Sequence 7897, Ap
3	508.8	30.3	1185	10	Sequence 7904, Ap
4	508.8	30.3	1185	10	Sequence 86, Appl
5	505.6	30.1	1230	10	Sequence 6265, Ap
6	503.6	30.0	1185	10	Sequence 9866, Ap
7	450.2	26.8	3309400	9	Sequence 6334, Ap
8	446.4	26.6	1188	9	Sequence 1, Appli
9	436	25.9	1200	9	Sequence 3437, Ap
10	436	25.9	1448	8	Sequence 7318, Ap
11	420.4	25.0	1185	10	Sequence 3, Appli
12	420.4	25.0	1185	10	Sequence 6962, Ap
13	413.8	24.6	1057	10	Sequence 6972, Ap
14	396.4	23.6	1197	10	Sequence 1309, Ap
15	394	23.4	15598	7	Sequence 9073, Ap
16	393.2	23.4	1185	10	Sequence 82, Appl
17	390.8	23.2	1182	10	Sequence 8344, Ap
18	378.6	22.5	1188	10	Sequence 4519, Ap
19	376.2	22.4	1185	10	Sequence 6385, Ap
					Sequence 3832, Ap

Sequence 575, App
Sequence 166, App
Sequence 164, App
Sequence 1, Appli
Sequence 154, App
Sequence 526, App
Sequence 88, Appl
Sequence 151, App
Sequence 171, App
Sequence 159, App
Sequence 1, Appli
Sequence 168, App
Sequence 456, App
Sequence 52, Appl
Sequence 155, App
Sequence 158, App
Sequence 147, App
Sequence 162, App
Sequence 160, App
Sequence 156, App
Sequence 152, App
Sequence 148, App
Sequence 18, Appl
Sequence 165, App
Sequence 157, App
Sequence 125, App

ALIGNMENTS

RESULT 1

US-09-815-242-7897
; Sequence 7897, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 7897
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1194)
US-09-815-242-7897

Query Match 31.4%; Score 528.4; DB 10; Length 1194;
Best Local Similarity 66.1%; Pred. No. 5.3e-133;
Matches 813; Conservative 0; Mismatches 381; Indels 36; Gaps 2;

QY 217 GCGGCGAGGGCAAGTTCGAGGCGACCAACACACACGTCAACATAGGACGACCATCGGCAT 276
Db 1 GTGGCTAAGGAAAAATTCGACGTAAACAAACCGCACGTCAACGTCGCGACCATCGGTAC 60

QY 277 GTCGACACCGGAAAGACCACTCTACCGCGCGCTGACCATGGTGTGCTCGCTCGCTCGGT 336
Db 61 GTTGACCATGCGAAGACCACTCTGACCGCTGCATGACCAAGGCTGCTCGCATACCTGG 120

QY 337 GCGAGCGGCGCTAAGAAGTACGACGAGATCGACGCGCCCGCGAGGAGCGCGCGGGGT 396
Db 121 GGTGGTTCGGCTCGTGTTCGATCAGATCGACAACGCGCGGAGAAAGCGCGCGGT 180

QY 397 ATCACCATCAACACCGCCACGTCGAGTACGAGACCGGACCGCGCTAGCGACAGTC 456
Db 181 ATCACCATCAACACCGCCACGTCGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 240

QY 457 GACTCGCGCGCGCGCGCGCTATGTCGAGTATGATGATGATGATGATGATGATGATGATG 516
Db 241 GACTCGCGCGCGCGCGCGCTATGTCGAGTATGATGATGATGATGATGATGATGATGATG 300

QY 517 GGTGCGATCTCGTGTATCGGTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 576
Db 301 GGTGCGATCTCGTGTATCGGTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360

QY 577 CTCCTCGCAACGAGTGGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 636
Db 361 CTCCTCGCAACGAGTGGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420

QY 637 CTCGACGAGGAGTCTCGAGTCTCGAGTCTCGAGTCTCGAGTCTCGAGTCTCGAGTCT 696
Db 421 CTCGACGAGGAGTCTCGAGTCTCGAGTCTCGAGTCTCGAGTCTCGAGTCTCGAGTCT 480

QY 697 TACGAGTACGAGGCGGAGCGTACCAATCTCGTGTGCTCGGCGCGCGCGCGCGCGCG 756
Db 481 TACGAGTACGAGGCGGAGCGTACCAATCTCGTGTGCTCGGCGCGCGCGCGCGCGCG 540

QY 757 GCTCTCATGGTCAACCTTCGTTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 816
Db 541 GCTCTCATGGTCAACCTTCGTTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 579

QY 817 TCGTTGGTTGAAGTGGATTCCTATATCCAGTCCGCGAGGCGGAGTCACTCCCG 876
Db 580 TCGTTGGTTGAAGTGGATTCCTATATCCAGTCCGCGAGGCGGAGTCACTCCCG 639

QY 877 TCTGCTGCTGTTGAAGTGGATTCCTATATCCAGTCCGCGAGGCGGAGTCACTCCCG 936
Db 640 TCTGCTGCTGTTGAAGTGGATTCCTATATCCAGTCCGCGAGGCGGAGTCACTCCCG 699

QY 937 CTATAGAGTGGCGACCGTCAAGATGGTGACACAGTTCGATATCGTTCGGAATCCGGG 996
Db 700 CTATAGAGTGGCGACCGTCAAGATGGTGACACAGTTCGATATCGTTCGGAATCCGGG 759

QY 997 ACCCGGAATCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056
Db 760 ACCCGGAATCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 819

QY 1057 GCGGAGACATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
Db 820 GCGGAGACATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879

QY 1117 ATGCTGCTGCGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1176
Db 880 ATGCTGCTGCGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 939

QY 1177 GTGCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1236
Db 940 GTGCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 999

QY 1237 TTCTACATCGCGACAACTGATGTGACAGGAGTGTGACTACGATTAATGATGACAAAGAT 1296
Db 1000 TTCTACTTCGTTACCACTACCGCATGTGACCGG-----TAACTGCGAAGTCCG 1044

QY 1297 GAGGAGCGAAGATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356
Db 1045 GAAGGCGTGAAGATGTAATGCGCGCGGACACATCAAGATGCTGCTGCTGCTGCTG 1104

QY 1357 CTTGTTGCTTGTGAGGAGGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1416
Db 1105 CCGATCGCATGGAAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1164

QY 1417 GCGCGTGTCTATCAACAAATCATTTAGTAA 1446
Db 1165 GCGCGGTGTTGCCAAGATCATCGAGTAA 1194

RESULT 2

US-09-815-242-7904
; Sequence 7904, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7904
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1194)
US-09-815-242-7904

Query Match 31.0%; Score 521.6; DB 10; Length 1194;
Best Local Similarity 65.8%; Pred. No. 3.6e-131;
Matches 808; Conservative 0; Mismatches 384; Indels 36; Gaps 2;

QY 219 GCGGAGGCGCAAGTTCGAGCGCACCAACACACGTCAACATAGGACCATCGGCATGT 278
Db 3 GCGTAAAGAAAATTTGAACGGAACAGCGCACGTCAACGTTCGGACCATCGGTACGT 62

QY 279 CGACCAAGGAAAGACACATCTTCCCGCGCGCTCACCATGCTCGCTCGCTCGCTCG 338
Db 63 TGACCATGGCAAGACCATCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122

QY 339 CAGCGCGCTAAGAGTACGACGAGTACGACCGCGCGCGCGCGCGCGCGCGCGGTAT 398

QY 757 GCTCTCATGGTCAACCCCTGCTTGAAGCGGGCGGACGATGAGTGGGTCGACTACATCTTC 816
Db 540 -----AGGCGACGAGTGGGAAGCGAAAATCCTG 570
QY 817 TCGTTGGTTGATAAAGTGGATTCCTATATTCAGTCCCGCAGAGCGACACTGACCTCCG 876
Db 571 GAACCTGGCTGCTTCCTGGATTCCTATATTCGGAACCGAGCGCTGGATTTGACAAGCCG 630
QY 877 TTCTTGCTGCTGTGTAAGATGCTTCTCCATACCCGCTGCTGTACAGTTGCCACTGGC 936
Db 631 TTCCTGCTGCCGATCGAAGAGCTATTCCTCATCTCCGCTGCTGTACCTGTGTACCGGT 690
QY 937 CGTATAGAGCTGGACCGCTCAAGATTTGGTACACAGTTCGATATTCGGAATCCCGGAC 996
Db 691 CGTGTAGAAGCCGCTATCATCAAGTTGGTGAAGAAGTTGAATCGTTGGTATCAAGAG 750
QY 997 ACCCGGAAGTCCAGGTCACCTGGTGTGAGATGTTCCAGAGACCATGGATGATGCCATG 1056
Db 751 ACTCAGAAGTCTACCTGTACTGGGCTTGAATGTTCCGCAACTGCTGGAGCAAGGCCGT 810
QY 1057 GCCGGAGACAATGTTGGGCTGCTGCTCGGTATGCAAGAGGATGACATTGAAAGAGC 1116
Db 811 GCTGCTGAGAACGTAGTGTCTGCTGCTGTATCAACGCTGAAGAAATCGACGTGGT 870
QY 1117 ATGGTCTGGCAAGCTGGCTCTATCACACCGCACCAAGTTTGAAGCTGTGTGTAT 1176
Db 871 CAGGTACTGGCTAAGCGGGGACCATCAAGCCGACACCAAGTTTGAAGTTCGAATCTGAAGTGTAC 930
QY 1177 GTGCTTAAGAGAGGAGGGTGGCGACATCACTTCTTCCCTGCTTACCGGCCACAG 1236
Db 931 ATTCGTCAAAGATGAAGCGCGCGCTCATCTCCGTTCTTCAAAGGCTACCGTCCCGAG 990
QY 1237 TTCTACATCGCGACAATGTTGACAGGGAGTGTGACTACGATTATGAATGACAAGGAT 1296
Db 991 TTCTACTTCCGCTACTGACGTGACTGG-----TACCATCGAAGTCCCG 1035
QY 1297 GAGGAGCGAAGATGTGCATCCCTGGTGACCGTATCAAAATGATTGTTCAGCTCATCCAG 1356
Db 1036 GAAGCGTAGAGATGGTAAATCCCGGGGACCAACATCAAAATGGTTGTACCTCGATCCAC 1095
QY 1357 CTTGTTGCTGTGACAGGGTATGAGTTTCTGCTATCCGCTGAGGTTGTTAAGACCGTTGGT 1416
Db 1096 CCGATCGGATGGACGAGGCTGCGTTTCCCAATCCGTGAAGCGCGCTACCGTTGGC 1155
QY 1417 GCCGCTGTCATCAACAAATCATTTAGTAA 1446
Db 1156 GCGGCGGTGTTGCTAAAGTTCTGGGCTAA 1185

RESULT 4

US-09-815-242-6265

; Sequence 6265, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6265
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1).....(1185)
US-09-815-242-6265

Query Match 30.3%; Score 508.8; DB 10; Length 1185;

Best Local Similarity 65.3%; Pred. No. 1.1e-127;

Matches 803; Conservative 0; Mismatches 382; Indels 45; Gaps 2;

QY 217 GCGGCGAGGGGCAAGTTCAGCGCACCAACACACAGTCAACATAGGCACCATCGCCAT 276
Db 1 GTGCTAAAGAAAAATTGAACGTACAAAACCGCAGTTAAGCTTGGTACTATCGGCCAC 60
QY 277 GTGACACACGGAAGAACCACTCTACCGCGCGCTCACCATGCTGCTCCCTCCGTCGGT 336
Db 61 GTTGACCAACGCTAAACTACTCTGACCGCTGCAATCACCCAGCTACTGGCTAAACCTAC 120
QY 337 GCGAGCGCCCTAAGAAGTACAGAGATCGAGCGGCGCCCGAGAGAGCGCCCGCGGT 396
Db 121 GCGGCTGCTGCTGCTGATTCGACAGATCGATAACGCGCCGGAAGAAAAGCTGTGCT 180
QY 397 ATCAACATCAACACCGCACCGTTCGAGTACGAGACGAGACCGCCACCTACGCACACGTC 456
Db 181 ATCAACATCAACACTTCTCACGTTGAATACAGACACCCCGCCGCTACTTACGCACAGTA 240
QY 457 GACTGCGCGCGCACCGCGACTATGTCAAGAAATATGATCACCGCGCTCGCGAGATGGAC 516
Db 241 GACTGCGCGCGCACCGCGACTATGTAAAAACATGATCACCGGCTGCTCAGATGGAC 300
QY 517 GGTGCCATCCTGCTGCTATCCGTCGCCGAGCGCCCATCGCCGACGACCAAGACACATC 576
Db 301 GCGCGCATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 577 CTCTCGCCCAAGCAAGTCCGCTTCCCAAGATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 636
Db 361 CTGCTGGGTCGTCAGTACGCTAGCGCTTCCGTATCATCTGCTTCTTCTTCTTCTTCTTCT 420
QY 637 GTCGACGACGAGAGTCTGCTGAGCTGCTGAGCTCGAGGTCGCGGAGCTGCTCAGCAAC 696
Db 421 GTTGATGACGAAGAGCTGCTGGAACCTGTTGAAATGGAAGTTCTGTAACCTTCTGCTCAG 480
QY 697 TAGGATGACGAGCGGACGACGCTACCAATCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 756
Db 481 TAGGATGCTCGGCGGACGACACTCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
QY 757 GCTCTCATGGTCAACCCCTGCCCTTGAAGCGCGGCGAGATGAGTGGTTCGACTACATCTTC 816
Db 540 -----AGGCGACGACAGTGGGAAGCGAAAATCCTG 570
QY 817 TCGTTGGTTGATAAAGTGGATTCCTATATTCAGTCCCGCAGAGCGACACTGACCTCCG 876
Db 571 GAACCTGGCTGCTTCCTGGATTCCTATATTCGGAACCGAGCGCTGGATTTGACAAGCCG 630
QY 877 TTCTTGCTGCTGTGTAAGATGCTTCTCCATCACCGGTCGTTGGTACAGTTGCCACTGGC 936
Db 631 TTCTGCTGCCGATCGAAGAGCTATTCCTCATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 690
QY 937 CGTATAGAGCTGGCACCGCTCAAGATTTGGTGAACAGTTCGATATTCGGAATCCCGGAC 996

Db	691	CGTGTAGACGCGGTATCATCAAGATTGGTGAGAAAGTTGAAATCGCTGGTGTATCAAGAG	750
Qy	997	ACCCGGAAGTCACGGTCACTGGTGTGAGATGTTTCCAGAAGACCATGGATGATGCCATG	1056
Db	751	ACTCAAAAGTCTACCTGTATTCGGCTTGAAATGTTCCGCAAACTGCTGGACGAAGCCGT	810
Qy	1057	GCCGGAGACAATCTTGGGCTGCTCCGTGGTATGCAAGAGGATGACATTGAAGAGGC	1116
Db	811	GCTGGTGAGAACGTAGGTGTTCTGCTGGGTGATATCAAACTGGAAGAAATCGAACGTGGT	870
Qy	1117	ATGCTGCTGCGAAAGCCTGGCTTCTATCACCGCACACCAAGTTTGAGGGTGTGTGTAT	1176
Db	871	CAGGTACTGGCTAAGCGGGCACCATCAAGCGGCACACCAAGTTCGAATCTGAAGTGTAC	930
Qy	1177	GTGCTTAAGNAGGAGAGGGTGGCGACACTCACTTCTTCTCCCTGGTTACCGGCCACAG	1236
Db	931	ATTCTGTCCAAAGATGAAGCGCGCTCATACTCGTGTCTTCAAAGGCTACCGTCCGAC	990
Qy	1237	TTCTACATCGGACAACTGATGTGACAGGGAGTGTGACTACGATTATGATGACAAGGAT	1296
Db	991	TTCTACTTCGGTACTGACGTGACTGG-----TACCATCGAAGTCCG	1035
Qy	1297	GAGGAGCGAAGATGTGCATGCCTGGTGACCGTATCAAAATGATGTTCAGCTCATCCAG	1356
Db	1036	GAAGCGGTAGATGTGTAATGCGGGCGCACAACTCAAAATGGTTGTTCCTGTGATCCAC	1095
Qy	1357	CGTGTGCTTGTGACAGGATGAGTTTGCATCCGTGAGGCTGGTGAAGCCGTTGGT	1416
Db	1096	CCGATCGCAATGGACAGCGGTCTCGTTTCGCAATCCGTGAAGCGGCCGTACCGTTGGC	1155
Qy	1417	GCCGGTGTTCATCAACAAATCATTAGTAA	1446
Db	1156	GCGGCGCTTGTGCTAAAGTCTTCGGCTAA	1185

RESULT 5

```

US-09-815-242-9866
: Sequence 9866, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011a
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9866
: LENGTH: 1230
: TYPE: DNA
: ORGANISM: Salmonella typhi

```


; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 3437
 ; LENGTH: 1188
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3437

Query Match 26.6%; Score 446.4; DB 9; Length 1188;
 Best Local Similarity 63.8%; Pred. No. 8.4e-111;
 Matches 788; Conservative 0; Mismatches 391; Indels 57; Gaps 5;

QY	217	GGGCGAGGGGCAAGTTCGAGCGACCAACACACAGGTCAACATAGGCACCATCGGCAT	276
DB	1	GTGGCAAGGGCAAGTTCGAGCGTACCAAGCCACGTAACATCGGCACCATCGGCAT	60
QY	277	GTCGACCGAAGAACACCTCTCACCGCGCGCTCACCATGGTGTCT-----CGCCTCC	330
DB	61	GTGACACCGGTAAAGACACCAACACCGCGCTATCACCAAGTTCGTGCTGACACTTAC	120
QY	331	GTCGTCGCGCGCCCTTAAAGATACAGAGATCGACGCGCGCGCGGAGGAGCGGCC	390
DB	121	CTGAGCTCAACAGAGGCTTCGCGCTTCGACTCCATCGATTAAGGCTCTCGAGGAAGGAG	180
QY	391	CGCGGTATCACATCAACACCGCGCGCTCGAGTACGAGACCGCGCGCTACGCA	450
DB	181	CGTGCATCACATCAACATCTCCACGTTGAGTACGAGTGAAGCGGCATACGCA	240
QY	451	CACGTGACTCGCGCGCGCGCGCTATGTCGAAGATATGATCACCGCGCTCGCGAG	510
DB	241	CACGTGACTCGCGCGCGCGCGCTATGTCGAAGATATGATCACCGCGCTCGCGAG	300
QY	511	ATGAGCGGTGCATCTCGTGTATCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	570
DB	301	ATGAGCGGTGCATCTCGTGTATCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	360
QY	571	CACATCTCTCGCGCGCGCGCGGTTCGCCAAGATCGTGTCTTCTTCTCAACAAGAG	630
DB	361	CACGTCTCTCTCGCGCGCGGTTCGCCAAGATCGTGTCTTCTTCTTCAACAAGAG	420
QY	631	GACATGTCGACGAGGAGCTGCTCGAGCTCGTGCAGTTCGAGTTCGCGAGCTGCTC	690
DB	421	GACATGTCGAGGATGAGGAATCATCGAGCTCGTGCAGTTCGAGTTCGAGTTCGCT	480
QY	691	AGCACTACGAGTACGAGCGCGCGCGCGCTACCAATCGTGCAGTTCGCGCGCTCAAGCG	750
DB	481	GCTGAGCGGACTACGA---CGAAGAGGCTTCAATTTTCACATCTCGCTCTGAAGGCT	537
QY	751	CTCAGGCTCTCATGGTCAACCCCTGCCTTGAAGCGCGCGCGCGCGCGCGCGCGCTAC	810
DB	538	CTTGAG-----GGCAGCAGAGAGTGGGCGAAGCAG	567
QY	811	ATCTTCGTTGGTTGATAAGTGGATTCCTATATTCAGTTCGCGCGCGCGCGCGCGCTGAC	870
DB	568	ATCTTCGAGTTCATGAGGCTTGGATGACAACTCCCTGACCGAGTTCGTCGAGCGGAC	627
QY	871	CTCCGCTCTGCTGCTGTTGAGATGCTCTTCCATCAGCGGTGCTGTCAGTTCGCG	930
DB	628	AAGCCATCTCTCATGGCTTCGAGGAGATCTTACCATCAGCGGTGCTGCGACCGTGT	687
QY	931	ACTGCGGTATAGAGCGTGGCGCGCTCAAGATTGGTGACACAGTCGATATCGTCGGAATC	990

DB	688	ACCGGTGCTGTTGAGCGCGGTACCCCTGAACGTTGAACGATGATGTTGACATATCGGCATC	747
QY	991	CGGACA---CCCGAACTGCACGGTCACTGGTGTGAGATGTTCCAGAAAGACCATGGAT	1047
DB	748	AAGGAGAGTCCACCTCCACCGCTTACCGGTATCGAGATGTTCCGTAAAGCTTCTTGAC	807
QY	1048	GATGCCATGCCGCGGAGACAATGTTGGGCTGCTGCTCCGTGGTATGAGAAAGATGACAT	1107
DB	808	TCCGCTGAGGCTGGCGACAACCTGCTGCTCTCTCCGTGGTATCAAGCGCGAAGATGTT	867
QY	1108	GAAAGAGGATGCTGCTGCGCAAGCGCTCTATCACACCGCACCAAGTTTGAGGCT	1167
DB	868	GAGGCTGGCAGGTTATCGTTAAGCCAGCGCTTACCCCTCACACCGAGTTCGAGGCG	927
QY	1168	GTTGTGTATGCTTAAAGAGGAGAGGTTGGCGCGACACTCACTTCTTCCCTGGTTAC	1227
DB	928	TCTGTCTAGTTCCTCAAGGATGAAGTGGCGCGCACACCCCATCTTCGACAACTAC	987
QY	1228	CGCCACAGTTCATATGCGGACAACATGATGTGACAGGAGTGTGACTACGATATGAAT	1287
DB	988	CGTCCTCAGTTCCTACTTCGCGCACCCAGCTTACCGGTGTTGTGA-----	1033
QY	1288	GACAAGGATGAGGAGCGGAGATGCTCATGCTGCGTGGTACCGTATCAAAATGATTGTCAG	1347
DB	1034	-AGTTCAGAGGCGACCGAGATGTCATGCTGGCGACAACGTCGATGTCGCTCACC	1092
QY	1348	CTCATCAGCGCTGCTGTTGTGAGCAGGATGATGAGTTCGCTATCGGTGAGGCTGTAAG	1407
DB	1093	CTGATCAGCGCTGCTGCTATGAGCAGGCGCTCGCTTTCGCTATCGCGAAGGCTCCCGC	1152
QY	1408	ACCGTTGGTGGCGGTCTCATCAACAAATCATTTGAG	1443
DB	1153	ACCGTTGGCGCTGCTGCTGTCACCAAGATCATCAAG	1188

RESULT 9
 US-09-815-242-7318
 ; Sequence 7318, Application US/09815242
 ; Patent No. US2002061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7318
 ; LENGTH: 1200
 ; TYPE: DNA
 ; ORGANISM: Helicobacter pylori

FEATURE:

NAME/KEY: CDS
LOCATION: (1)...(1200)
US-09-815-242-7318

Query Match 25.9%; Score 436; DB 10; Length 1200;
Best Local Similarity 61.3%; Pred. No. 5.6e-108;
Matches 753; Conservative 0; Mismatches 445; Indels 30; Gaps 2;

QY 219 GCGAGGCGCAAGTTTCGAGCGCACCAAAACACACAGTCAACATAGGACCAATCGGCCATGT 278
Db 3 GCGAAAGAAAGTTTAAACAGAACTAAGCGCATGTTAATATTGAAACCAATGGGCATGT 62
QY 279 CGACACCGAAGAACCACTCTCACCGCGCGCTCACCATGCTGCTCGCCCTCCGTCGGTGG 338
Db 63 AGACCATGTTAAAGAGCTTTGAGTGCAGCGATTTCAGCGGTGCTTTCTTTGAAAGTCT 122
QY 339 CAGCGCGCTAAGAGTAGACGAGATCGACGCGCGCCCGCCGAGGCGCGCGCGGTAT 398
Db 123 TCGAATATGAAGACTATGATATATTGATAACGCCCTGAAGAAAAGAGGGAT 182
QY 399 CACATACACCGCGCCCGCTGAGTAGACGAGCGAGACCGCCCACTACGACACGTCGA 458
Db 183 CACTATCTACTTCTCATATGATATGAGACTGAAACACAGACACTATGCGCATGTGA 242
QY 459 CTGCGCGCGCACGCGGACTATGTCAAGAATATGATACCGCGCGCTGCGCAGATGGACGG 518
Db 243 TTGCCCGAGACAGCTGACTATGTAAACACATGATACCGCGCGCGCAATGGACGG 302
QY 519 TGCCATCTCTGCTGCTATCGGTGCGGCGCGCCATGCGGAGACCAAGAGACACATCT 578
Db 303 AGCGATTTTGGTTGTTCTGCGAGCTGATGGCCCTATGCTCAAACTAGGAGCATATCT 362
QY 579 CCTCCCAAGACAGTGGTGTCCCAAGATGTTGCTTCTCTCAACAAGAGGACATGGT 638
Db 363 ATTGCTCTGCTCAAGTAGGCGTGCCTCATCATGTTGTTTCTTAAACAACAAGACATGGT 422
QY 639 CGACACGAGGAGCTGCTCGAGCTGCTGAGCTCGAGTCCGCGAGCTGCTCAGCACTA 698
Db 423 AGATACCAAGAAATTTGAGAACTGTGAGAAATGGAAGTGGCGGAATTTGAGCGCGTA 482
QY 699 CGAGTACGACGCGGACGACGCTACCAATCGTCCGCTGCGCCCTCAAGGCGCTCGAGGC 758
Db 483 TGAATTTCTGCGGATGACACTCTCTATCTGAGCGGTTTCAGCTTTAAGCTTTAGAAG- 541
QY 759 TCTCATGTCACCTGCTTGAACGCGGCGGACCATGAGTGGTCTGCTACATCTCTC 818
Db 542 -----AAGCAAGGCTGGTAATGTTGGGTGAATGGGGTGAAGGCTGCTTAA 587
QY 819 GTTGGTTGATAAGTGGATTCCTATATTCAGTCCCGCAGAGGAGACTGACCTCCCGTT 878
Db 588 ACTTATGCTGAGTGGATGCTATATCTCTACTCCAGAAAGAGACACTGAAACACTTT 647
QY 879 CTTGCTGCTGTTGAAGATGCTTCTCCATCACCGGTGCTGCTGAGTGGCCACTGGCGG 938
Db 648 CTTGATGCGGTTGAAGATGCTTCTCTATTGCGGTTAGAGGACTGTTGTTACAGTAG 707
QY 939 TATAGAGCTGCGACCGCTCAAGATGTTGACACAGTCCATATCGTCGGAATCCGGACAC 998
Db 708 GATTGAAGAGCGGTGGTGAAGTAGGCGATGAAGTGAATCGTTGTTATCAGACCTAC 767
QY 999 CCGGAAGTGCAGGCTGCTGTTGAGATGTTCCAGAGAGCAATGATGATGCTATGCC 1058
Db 768 ACAAAAAGCACTGTAAACCGGTGTAGAAATGTTAGGAAAGAGTTGGAAGAGGTGAAGC 827
QY 1059 CGGAGCAATATGTTGGGCTGCTGCTCCGCTGATGAGAGAGGATGACATTTGAAAGGCAAT 1118
Db 828 CGCGATATATGTTGGGCTGCTTTCAGAGGAACTAAAAAAGAAAGTGAAGCGCGTAT 887
QY 1119 GGTGCTGCAAGGCTGCTCTATCACACCGCACCAAGTTTGGGCTGTTGCTATGT 1178
Db 888 GGTCTCTGCAAAACAGGTTCTCATCTCCGACCAAGAAATTTGAGGAGAAATTTATGT 947

QY 1179 GCTTAAGAGGAAGAGGTGGCGGACACACCTCTTCTCCCTGGTTACGCCACAGTT 1238
Db 948 CCTTCTTAAAGAAAGAGCGGAGACACACTCCATCTCTTCCACCAATACGCCCGCAAT 1007
QY 1239 CTACATCGCGGACAACTGATGTGACAGGAGTGTGACTACGATTTATGAATGACAAGATGA 1298
Db 1008 CTATGTGCGCACAACTGATGTGACTGGCTCTATCAC-----CCTTCTCTGA 1052
QY 1299 GGAGGCGAAGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1358
Db 1053 AGCGTAGAAATGTTTATGCTGCGGATATGTGAAATACACTGTAGAGTTGATTAGCCCC 1112
QY 1359 TGTGCTGTGAGCAGGATGAGTGTGCTATCCGCTGAGGCTGCTGAGGCGGTAGGACCGTGGTGC 1418
Db 1113 TGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1172
QY 1419 CGGTGCTATCAACAAATCATTTAGTAA 1446
Db 1173 TGTGTTGTGAGCAATATTATTGAATAA 1200

RESULT 10

US-08-831-310-3
Sequence 3, Application US/08831310
Patent No. US2002026035A1
GENERAL INFORMATION:
APPLICANT: Kleanthous, Harold et al.
TITLE OF INVENTION: Helicobacter GHPO 1360 and
TITLE OF INVENTION: GHPO 750 Polypeptides and Corresponding Polynucleotides
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/831,310
FILING DATE: 01-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,175
REFERENCE/DOCKET NUMBER: 06132/037001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1448 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 118..1314
OTHER INFORMATION:
US-08-831-310-3

Query Match 25.9%; Score 436; DB 8; Length 1448;
Best Local Similarity 61.3%; Pred. No. 6.1e-108;
Matches 753; Conservative 0; Mismatches 445; Indels 30; Gaps 2;
QY 219 GCGAGGCGCAAGTTTCGAGCGCACCAAAACACACAGTCAACATAGGACCAATCGGCCATGT 278
Db 1119 GGTGCTGCAAGGCTGCTCTATCACACCGCACCAAGTTTGGGCTGTTGCTATGT 1178
Db 888 GGTCTCTGCAAAACAGGTTCTCATCTCCGACCAAGAAATTTGAGGAGAAATTTATGT 947

Db 120 GCACAAAGAAAGTTTAAACAGAACTAAGCCGATGTTAAATATGGAACCATTTGGGCATGT 179
Qy 279 CGACACGAAAGACCACTCTACCGCGCGCTCACCATGGTCTCGCTCCGCTCGGTGG 338
Db 180 AGACCATGTAAGACACTTTAGTGCAGCGATTTTACGGGTGTTCTTTGAAGGTCT 239
Qy 339 CAGCGCGCTTAAAGATACGACGATGCGCGCGCGCGCGCGCGCGCGCGCGGTAT 398
Db 240 TGCAGAAATGAAGACTATGATAATATTGATAAGCCCTCAAGAAAAAAGAGAGGAT 299
Qy 399 CACCATCAACCGCCACCGCTAGTACGAGACCGAGACCGCCACTACGCACACGTCGA 458
Db 300 CACTATCCCTACTTCTCATATTGAATATGAGACTGAAACAGACACTATGCGATGGGA 359
Qy 459 CTGCCCGCCACCGCACTATGTCAGAAATATGATCACCGCGCTCGCGAGATGGACGG 518
Db 360 TTGCCAGACACGCTGACTATGTAAAAACATGATCACCGGTGCGGCGCAATTTGACGG 419
Qy 519 TGCATCTCTGATTCGGTGGCGGACGGGCGCATCCGCGAGACCAAGAGACATCCT 578
Db 420 AGCGATTTGGTTGTTCTGACGCTGATGGCCCTATGCCTCAAACTAGGAGCATATCTT 479
Qy 579 CTTGCCCAAGCAAGTCCGTGTTCCAGATCGTTGTTCTTCTCAACAAGAGACATGTT 638
Db 480 ATTGTCTCGTCAAGTAGCGGTGCTCATCTGTTGTTTCTTAAACAAACAGACATGTT 539
Qy 639 CGAGCAGGAGGAGTGTCTGAGCTGCTGAGCTCGGAGTCCGCGAGCTGCTCAGCAACTA 698
Db 540 AGATGACCAAGAATTTAGAACTTGTAGAAATGGAAGTGGCAATTTGTAGCGGCTA 599
Qy 699 CGAGTACAGCGGACGACGCTACCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 758
Db 600 TGAATTTCTCGCATGACACTCTCTATCTGCTAGCGGTTGAGCTTAAAGAGCTTTAGAAG - 658
Qy 759 TCTCATGTCAACCTCCCTTGAGCGCGGAGATGAGTGGTGGCTACATCTTCTC 818
Db 659 -----AAGAAAGGCTGTAATGTTGGTGAATGGGTGAAAGAGTCTTAA 704
Qy 819 GTTGGTGTAAAGTGGATTCCTATATCCAGTCCCGCAGAGGACAGCTGACCTCCGTT 878
Db 705 ACTTATGGCTGAAGTGGATGCTATATCCCTACTCCAGAAAGAGACACTGAAAGACTTT 764
Qy 879 CTTGCTCGCTTCAAGATGCTCTCCATCAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 938
Db 765 CTTGATCGCGTGAAGATGTTCTCTATTTGCGGCTAGAGGAGTGTGTTACAGGTAG 824
Qy 939 TATAGAGCTGGCCCGCTCAAGATTTGCTGACAGCTGATATCGTGGAAATCCGGACAC 998
Db 825 GATTGAAGAGCGCTGTGAAGTAGCGGATGAAGTGGAAATCGTTGTTATCAGACCTAC 884
Qy 999 CCGAACTGCACGCTCACTGGTGTGATGTTCCAGAGACCATGATGATGCCATGGC 1058
Db 885 ACAAAAACAGCTTAACCGGTGTAGAAATGTTTAGGAAAGAGTTGAAAGAGTGAAGC 944
Qy 1059 CGGAGCAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
Db 945 CGGCGATAATTTGGCGCTGCTTTTGAGAGAACTAAAGAAAGAGAGTGGAAACGCGGTAT 1004
Qy 1119 GGTGCTGGCAAGGCTGGCTTATCACACCGCACACCAAGTTTGAAGCTGTTGTATGT 1178
Db 1005 GGTCTATGCAACACAGTTCTATCTACCTCCGACAGAAATTTGAGGAGAAATTTATGT 1064
Qy 1179 GCTTAAAGAGAGAGGTTGGCGACACTCACCTTTTCCCTTGGTTACCGCCACAGTT 1238
Db 1065 CTTTCTTAAAGAGAGGCGGAGACACACTCTCATTTCTTCAACCAATTTACCGCCGCAAT 1124
Qy 1239 CTACATCGGACACTGATGTGACGAGGAGTGTGACTACGATTATGAATCACAAGATGA 1298
Db 1125 CTATGCGGCACACTGATGTGACTGGCTCTATCAC-----CCTTCTGA 1169
Qy 1299 GGAGGCGAAGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
Db 1170 AGCGGTAGAAATGTTATGCTGGCGATATGTGAAATCACTGTAGATTGATTAGCCC 1229

Qy 1359 TGTTCCTTGTGACAGGCTATGAGTTTGTCTATCCGTGAGGTGGTAAGACCGTTGGTGC 1418
Db 1230 TGTTCCTTGTGAGTTGGGAACTAAATTTCCGATTCGTGAAGCGGTAGGACCGTTGGTGC 1289
Qy 1419 CGTGTGTCATCAACAAATCATTTAGTAA 1446
Db 1290 TGTGTGTGTGAGCAATATTATTGAATAA 1317

RESULT 11

US-09-815-242-6962
; Sequence 6962, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6962
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1185)
US-09-815-242-6962

Query Match 25.0%; Score 420.4; DB 10; Length 1185;
Best Local Similarity 60.8%; Pred. No. 9.3e-104;
Matches 747; Conservative 0; Mismatches 436; Indels 45; Gaps 2;

Qy 219 GCGAGGGGCGAGTTCGAGGCGCACCAACACACAGCTCAACATAGCACCATCGCCATGT 278
Db 3 GTCTAAAGAAAATTTGAACGTACAAAACCGCACCTAAACGTGGTACATTCGGCCACGT 62
Qy 279 GCACACGGAAGACACTCTCACCGCGCTCACCATGTTGCTTCGCTCCGTCGGTGG 338
Db 63 TGACACCGTAAACAACTTTAAACAGCAGCAATCAACACCGTATTAGCAAAACACTAGG 122
Qy 339 CAGCGCGCTTAAAGATGACGAGATCGACCGCGCGCGCGCGCGCGCGCGCGGTAT 398
Db 123 TGTGTACGCGGTGCTATTGACCAAAATTTGTAACGCGCGCAGAAAAGCGCGGTAT 182
Qy 399 CACCATCAACCGCGCGCTGAGTACGAGACCGAGACCGCCGCTACGACACAGTCA 458
Db 183 TACCATCAACACTTCACAGCTTGAATAGTACACCGACTCGCCACTACGACACAGTAGA 242


```

QY 459 CTGCCCCGCGCCAGCGGCTATGTCACGAATATGATCACCGCGCTGCGCAGATGGAGG 518
Db 243 CTCTCGGAGACGCGGCTATGTTAAATATGATTTACTGTGCGGCACAAATGGATGG 302
QY 519 TGCATCTCTGCGGTATCGGTGCGGACGCGGCCCTGCGGACGACCAAGAGACATCCT 578
Db 303 TGCATTTTGTAGTAGTAGCAGCAACAGATGGTCTTATGCGACAACTCTGTAACACATCT 362
QY 579 CTCTGCCAAGCAAGTCGGTGTTCACCAAGATCGTGTCTCTCTCTCAACAAGAGGACATGGT 638
Db 363 ATTAGTCCCAAGTAGGTGTTCATATCATCTGATTTCTTAAACAAATCGCATGGT 422
QY 639 CGACGACGAGGAGCTCTCGAGCTCGTGCAGCTCGAGTCCGCGAGCTGCTCAGCAACTA 698
Db 423 AGATGACGAAGAGTTATTAGATTTAGTCGAATGGAAGTTCGTGAATCTTATCTCAATA 482
QY 699 CAGATACGACGCGGAGGAGTACCAATCGTCTGCTGCTCGGCTCAAGGCGCTCGAGGC 758
Db 483 TGACTTCCGAGGTGAGCATACCAATCGTACGTGGTTTCAGCATTTACAAGCGTTAAA--- 539
QY 759 TCTCATGTCTAACCCCTGCTTGAAGCGCGGACGATGAGTGGTTCGACTACATCTTCTC 818
Db 540 -----CGGCGTAGCAGAAATGGGAGAAAATCTCTGA 572
QY 819 GTTGTGTGATAAGTGGATTCCTATATTCAGTCCCGCAGAGGAGACTGACCTCCCGTT 878
Db 573 GTTAGCAACCACTTAGATCTTACATCCAGAACGAGACGTCGATTTGACCAACCGTT 632
QY 879 CTGCTCTGCTGTGAAGATGCTTCTCTCATCACCGGTGCTGAGTGGTCCATGCGCG 938
Db 633 CTTCTTCCATCGAAGATGTTCTCAATCTCAGGTGCTGTACTGTAGTAACAGTCTG 692
QY 939 TATAGAGGTGCGACCGCTCAAGATTTGGTGCACACAGTCTGATCTCGGAATCCGGGAC 998
Db 693 TGTAGAACGAGGTATTATCCGTACAGGTGATGAAGTAGAATCTCGGTATCAAGATAC 752
QY 999 CGGAACTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1058
Db 753 AGCGAAACTACTGTAAAGGTTGTAATGTTCCGTAATTTACTTGACGAAGTCTGTC 812
QY 1059 CGGACCAATGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
Db 813 AGGTGAACCATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 872
QY 1119 GGTGTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1178
Db 873 AGTATTAGCAAAACAGGTTCATACACACACACACTGACTTCCGAATCAGAAAGTGTAGT 932
QY 1179 GCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1238
Db 933 ATTATCAAAAGATGAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 992
QY 1239 CTACATCGGACAACTGATGTGACAGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
Db 993 CTATTCGTACACAGAGTACTG-----TACAATCGAATTAACAGA 1037
QY 1299 GGAGCGAAGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
Db 1038 AGGCGTGAATGTTAATGCCAGGCGATAACATCAAGATGACAGTAAAGCTTAATCCACC 1097
QY 1359 TGTGTCTGTGACAGAGGTATGAGTTTGTATCCGTGAGGTTGTTAAGACCGTTGTTGTC 1418
Db 1098 AATTGCGATGATCAAGGTTTACGTTTTCGCAATCCGTTGAAGGTGGCGGTACAGTAGTGC 1157
QY 1419 CGGTGCTCATCAACAAATCATTTAGTAA 1446
Db 1158 AGGCGTGTTCGGAATCATCAATAAA 1185

```

RESULT 12

us-09-815-242-6972

; Sequence 6972, Application US/09815242

; Patent No. US20020061569A1

```

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6972
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; NAME/KEY: CDS
; LOCATION: (1)...(1185)
; US-09-815-242-6972

```

```

Query Match 25.0%; Score 420.4; DB 10; Length 1185;
Best Local Similarity 60.8%; Pred. No. 9.3e-104;
Matches 747; Conservative 0; Mismatches 436; Indels 45; Gaps 2;

```

```

QY 219 GCGAGGCGGCAAGTTCGAGCGCACCAACACACAGTCAACATAGGACACCATCGCCATGT 278
Db 3 GTCTAAGAAATTTGAAGGTACAAACCGCACGTAAACGTGGGTACAAATCGGCCACGT 62
QY 279 CGACACGGAAGAACCACTTCTACCGCGGCTCACCATCGCTCGCTCGCTCGCTCGGTGG 338
Db 63 TGACCACGCTAAACAACTTTAACAGCAGCAATCAACAACGTTATGGCAAAACATTCAGG 122
QY 339 CAGCGCGCTTAAGAAGTACGAGATCGAGCGCGCGCGCGCGCGCGCGCGCGCGGTAT 398
Db 123 TGTGCGGCGGTGTCATTCGACCAATTTGATAACGCGCGCAAGAAAGCGCGGTGTAT 182
QY 399 CACCATCAACACCGCGCTCGAGTACGAGACCGCGCGCGCGCGCGCGCGCGCGCGTCA 458
Db 183 TACCATCAACACTTCACAGTTGAATACGATACCGCTCGCCATATGCAACAGTAGA 242
QY 459 CTGCGCGCGCGCGCGCGCTATGTCAAGAAATATGATCACCGCGCTGCGCAGATGACGG 518
Db 243 CTGTCGGGACACGCGCGCTATGTTAAAAATATGATTACTGTTGGGCGCAAAATGGATGG 302
QY 519 TGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 578
Db 303 TGTATTTTGTAGTAGTAGCAGCAACAGATGCTCTATGCGCGCAAACTCTGTTGAACACATCT 362
QY 579 CTTGCGCAAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
Db 363 ATTAGTCCCAAGTAGGTGTTCCATCATCATCATCATCATCATCATCATCATCATCATCAT 422
QY 639 CGACGACGAGGAGTCTGCTGAGCTGCTGAGCTGCTGAGGTCCGCGAGCTGCTCAGCAACTA 698

```


GENERAL INFORMATION:

APPLICANT: Charles Kunsch
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5255
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Ge
 STREET: 9410 Key We
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM
 MEDIUM TYPE: Disket
 COMPUTER: HP Vectra
 OPERATING SYSTEM: M
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DMI
 APPLICATION NUMBER:
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION
 NAME: Benson, Bob
 REGISTRATION NUMBER: 307446
 REFERENCE/DOCKET NUMBER: PB248PP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 82:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15598 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

US-08-781-986A-82

Query Match 23.4%; Score 394; DB 7; Length 15598;

Best Local Similarity 57.6%; Pred. No. 4.2e-96;

Matches 774; Conservative 0; Mismatches 525; Indels 45; Gaps 2;

Qy 219 GCGAGGGGCAAGTTCGAGCGCCACCAACACACAGCTCAACATAGGCACCATCGGCCATGT 278
 Db 11537 GCGAAAAGAAAAATCGATCGTCTTAAGAACATATCGCTACTATCGGTACATCGGT 11596
 Qy 279 CGACCAGGAAGACCACTCTCACCGCGCGCTCACCATCGTCTCGCCCTCCGCTGG 338
 Db 11597 TGACCATGGTAAACAACATTAACAGCAGCAATCGCTACTGTATTAGCAAAAAATGGTGA 11656
 Qy 339 CAGCGCGCCTAAGAAGTACGACGAGATCGACGCGCGCCGCGAGAGCGCGCGCGGTAT 398
 Db 11657 CTCAGTTGCACAATCATATGACATGATTGACAGCTCCAGAGAAAAGAACGTGGTAT 11716
 Qy 399 CACCATCAACCCGACCGTCGAGTACGAGACGAGAGCCGCGCACTACGCCACAGTCGA 458
 Db 11717 CACAATCAATCTCTCACATTTAGTACCAAACTGACAAACGCTACGCTCACGCTGA 11776
 Qy 459 CTGCGCGCGCCGCGCACTATGTCAGAAATATGATCACCGCGCTCGCGAGATGGACGG 518
 Db 11777 CTGCGCAGGACACGCTGACTACGTAAACATGATGATCCTGGTCTCAATGACCG 11836
 Qy 519 TGCCATCCTCGTGTATCCGTCGCGCGCCGCGCCATGCCGAGACCAAGAGACATCCT 578
 Db 11837 CGGTATCTTAGTATCTGTGTCGAGCGTCCAAATGCCCAAACTCGTGAACACATCT 11896
 Qy 579 CTGCGCAAGCAAGTCGGTGTCCAGATCGTGTCTTCTCAACAGAGGACATGGT 638
 Db 11897 TTTATCAGCTAACGTTGGTGTACCAAGCATTAGTAGTATCTTAAACAAAAGTTGACATGGT 11956
 Qy 639 CGACGACGAGGCTGCTCGAGCTCGTCGAGCTCGAGCTCGGAGCTCGGAGCTGCTCAGCAACTA 698
 Db 11957 TGACGTGAAGAAATTTAGAAATTTAGTAGAATGGAAGATTCGAGTCTTATTAACCGAATA 12016

Qy 699 CGAGTACGACGGGACGACGACGTACCAATCGTCTGGCTCGCCCTCAAGCGCTCGAGGC 758
 Db 12017 TGACTTCCAGGTGACGATGTACCTGTAATCGCTGGTTCAGCATTAAGGCTTTAGA --- 12073
 Qy 759 TCTCATGGTCAACCCCTGCTTGAAGCGGGGCGAGATGAGTGGTTCGACTACATCTCTC 818
 Db 12074 -----AGCGATGCTCAATACGAAGAAAAATCTTAGA 12106
 Qy 819 GTTGGTGCATAAAGTGGATTCCTATATTCAGTCCCGAGAGCGAGACTGACCTCCCGTT 878
 Db 12107 ATTAATGGAAGCTGTAGATACTTACATTCACACTCCAGAGCTGATTTCTGCAACACATT 12166
 Qy 879 CTTGCTCGCTGTGAAGATGTCTTCCATCACCGGTCTGGTGTACAGTGCACACTGCGCG 938
 Db 12167 CATGATGCCAGTTGAGGACGTATCTCAATCACTGGTCTGCTGCTACAGCGCG 12226
 Qy 939 TATAGAGCTGGCACCGTCAAGATGGTGACAGTGCATATCTCGGAATCCGGGACAC 998
 Db 12227 TGTTCAGCTGGTCAAAATCAAAGTTGGTGAAGAAGTTGAAATCATCGGTTTACATGAC 12286
 Qy 999 CCGGAACCTGCACGCTCACTGGTGTGAGATGTTCCAGAAAGACCATGGATGATGCCATGGC 1058
 Db 12287 ATCTAAAACAACTGTTACAGGTGTGAATGTTCGGTAAATATTAGACTACGCTGAAGC 12346
 Qy 1059 CGGAGACAATGTGGGCTGCTGCTCGGTGTATGCGAAGAGATGACATTTGAAAGAGGAT 1118
 Db 12347 TGGTGACAAACATTTGGTGCATATTACGTGGTGTGCTCGTGAAGACGTACAACGTGTCA 12406
 Qy 1119 GGTGCTGCAAAAGCTGCTCTATCACACCGCACACCAAGTTTGAGGCTGTGTGTATGT 1178
 Db 12407 AGTATTAGCTGCTCTGTTCAATTACACCACTACTCAATTTCAAGCAGAGTATACGT 12466
 Qy 1179 GCTTAAGAAAGAGGGTGGCGGACACTCACCTTTCTTCCTGGTTACCGCCACAGTT 1238
 Db 12467 ATTATCAAAAGACGAAGTGGAGCTCACACTCCATCTCTCTCAAACTATCGTCCACAAT 12526
 Qy 1239 CTACATGGGCAACTGATGTGACAGGAGTGTGACTACGATTATGATGACAAGGATGA 1298
 Db 12527 CTATTTCCGTACTACTGACGTAACCTGGTGTGT-----TCACCTTACCAGA 12571
 Qy 1299 GGAGCGGAAGATGTCATGCTCGTGGTGACCGTATCAAAATGATTTGTACGCTCATCCAGCC 1358
 Db 12572 AGGTACTGAAATGGTAAATGCCCTGGTGTAAACGTTGAAATGACAGTAGAATTAATCGCTCC 12631
 Qy 1359 TGTTCCTTGTGAGCGGTATGAGTTTGTCTATCCGCGAGGTTGGTGAAGACCGTTGGTGC 1418
 Db 12632 ATTCGGATTGAAGACGGTACTCGTTTCTCAATCCGTGAAGGTGACGTACTGTAGGATC 12691
 Qy 1419 CGGTGTCATCAACAAAATCAATTGAGTAACTGGATATAACATATCCACCATGAGAAATTT 1478
 Db 12692 AGGCGTTGTTACTGAAATCAATTAATAATTTCTAATTTCTTAGATTTTATATAAAAGAA 12751
 Qy 1479 CTTTGTCTTACTCAAGCGACATGCTCCGTAGTTGTTATTATGTGGTGGTGTAGGGGTT 1538
 Db 12752 GATCCCTCAATCGAGGGGCTTTTTTTTAAATGTTGTAATTTTGTAAATGCTATTCGATT 12811
 Qy 1539 GCTCATGTGCAATTTGATGATGAC 1562
 Db 12812 GAAGAACAATAATTGATGAAGAC 12835

Search completed: May 1, 2003, 22:09:46
 Job time : 1226 secs